

**UTILITY PATENT APPLICATION  
TRANSMITTAL UNDER 37 CFR 1.53(b)**

**ATTORNEY DOCKET 79676DMW**

A

To: Commissioner for Patents  
Box Patent Application  
Washington, D.C. 20231

Express Mail Label No.

**EL485195513US**

Date: 3/29/00

jc682 U.S. PTO  
09/537334  
03/29/00

**A METHOD OF DETECTING DUPLICATE  
PICTURES IN AN AUTOMATIC ALBUMING  
SYSTEM**

First Named Inventor (or Application Identifier):

Alexander C. Loui, et al

Enclosed are:

1. ☒ Specification
2. ☐ 5 Sheet(s) of drawing(s)
3. ☒ Information Disclosure Statement Under 37 CFR 1.97.
4. Combined Declaration for Patent Application and Power of Attorney:
  - 4a. ☐ New
  - 4b. ☐ Copy from a prior application (37 CFR 1.63(d) (for continuation/divisional with Box 11 completed)
5. ☐ Incorporation by Reference (useable if Box 4b is checked) The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 4b, is considered as being part of the disclosure of the accompanying application and is hereby incorporated by reference therein.
6. ☒ Assignment of the invention to **Eastman Kodak Company**
7. ☐ Certified copy of a priority document
8. ☒ Letter under Rule 53
9. ☐ Deletion of Inventor(s). Signed statement attached deleting inventor(s) named in the prior application, see 37 CFR 1.63(d)(2) and 1.33(b).

10. ☐ If a 111A application prior to examination of the above-identified application, amend the specification at Page 1, after the title, by inserting the following:  
--CROSS REFERENCE TO RELATED APPLICATION  
Reference is made to and priority claimed from U.S. Provisional Application Serial No. , filed , entitled .

**If a CONTINUING APPLICATION**, check appropriate box and supply the requisite information:

11. ☐ Continuation ☐ Divisional ☐ Continuation-in-part (CIP) of prior application No: ,
12. ☒ Please address all written communications to Thomas H. Close, Patent Legal Staff, Eastman Kodak Company, 343 State Street, Rochester, NY 14650-2201.  
Please Direct all telephone calls to David M. Woods at (716) 477-5256.

The filing fee has been calculated as shown below:

FOR:	NO. FILED	NO. EXTRA	RATE	FEE
BASIC FEE				\$ 690
TOTAL CLAIMS	24	- 20 =	4	x 18 = \$ 72
INDEPENDENT CLAIMS	4	- 3 =	1	x 78 = \$ 78
<input type="checkbox"/> MULTIPLE DEPENDENT CLAIM PRESENTED			+ 260	\$0
			<b>TOTAL</b>	<b>\$ 840</b>

- ☒ Please charge my Eastman Kodak Company Deposit Account No. **05-0225** in the amount of \$ 840.

**A duplicate copy of this sheet is enclosed**

- ☒ The Commissioner is hereby authorized to charge any additional filing fees required under 37 CFR 1.16 or credit any overpayment to Eastman Kodak Company Deposit Account No. **05-0225**.

**A duplicate copy of this sheet is enclosed.**

DMW/RGR

Telephone: (716) 477-5256  
Facsimile: (716) 477-4646

David M. Woods  
David M. Woods  
Attorney for Applicants  
Registration No. 27,171

PATENT APPLICATION BASED ON: Docket Number 79676DMW

Inventor(s): Alexander C. Loui  
Eric S. Pavie

Attorney: David M. Woods

Document ID: 79676DMW\DOCKETS\

**A METHOD OF DETECTING DUPLICATE PICTURES IN AN  
AUTOMATIC ALBUMING SYSTEM**

EASTMAN KODAK COMPANY

RESTRICTED INFORMATION

"Express Mail" mailing label number EL485195513 US  
Date of Deposit March 29, 2000

I hereby certify that this paper or fee is being deposited  
with the United States Postal Service "Express Mail Post  
Office to "Addressee" service under 37 CFR 1.10 on the date  
indicated above and is addressed to the Commissioner of  
Patents and Trademarks, Washington, D.C. 20231

Robin G. Reeves

(Typed or printed name of  
person mailing paper of fee)

Robin G. Reeves

(Signature of person mailing paper or fee)

**A METHOD OF DETECTING DUPLICATE PICTURES IN AN  
AUTOMATIC ALBUMING SYSTEM**

**CROSS-REFERENCE TO RELATED APPLICATION(S)**

5           The present application is related to U.S. Application Serial  
Number 09/163,618, filed September 30, 1998, by Alexander C. Loui and Eric S.  
Pavie, and entitled, "A METHOD FOR AUTOMATICALLY CLASSIFYING  
IMAGES INTO EVENTS" and to U.S. Application Serial Number 09/197,363,  
filed November 20, 1998, by Alexander C. Loui and Eric S. Pavie, and entitled,  
10   "A METHOD FOR AUTOMATICALLY COMPARING CONTENT OF  
IMAGES FOR CLASSIFICATION INTO EVENTS".

**FIELD OF THE INVENTION**

15           The invention relates generally to the field of image processing  
systems that automatically classify pictures by events and the like and, more  
particularly, to an automatic classification and albing system that automatically  
classifies pictures for placement into an album.

**BACKGROUND OF THE INVENTION**

20           Pictorial images are often classified by the particular event,  
subject, or the like for convenience of retrieving, reviewing, and albing of the  
images. This classification is usually achieved by either manually or  
automatically segmenting the images into appropriate groups. A manual method  
would involve visually inspecting each image and then placing the image into the  
25   appropriate group. An automated albing method would typically group the  
images in some automatic manner by color, shape or texture in order to partition  
the images into groups of similar image characteristics.

          Although the presently known and utilized methods for  
partitioning images are satisfactory, there are drawbacks. The manual  
30   classification method is obviously time consuming, and the automated albing  
method, although theoretically classifying the images into events, is susceptible to  
misclassification due to the inherent inaccuracies involved with classification by  
color, shape or texture. In either method, when two pictures are identified as

duplicates, one of them is typically removed and will not appear in the resulting album. Since a consumer will not be satisfied if the automatic albuming process removes a picture that should have been in the album, the precision of the duplicate detection algorithm has to be high. Consequently, a need exists for  
5 overcoming the above-described drawbacks.

### SUMMARY OF THE INVENTION

The present invention is directed to overcoming one or more of the problems set forth above. Briefly summarized, according to one aspect of the  
10 present invention, the invention resides in a method for detecting duplicate images comprising the steps of providing at least two images captured at determinable times; computing an indication of the image content for each image; determining the time of capture of each of the images; and evaluating the indication of image content and the time of capture to determine whether the images are duplicate  
15 images.

In a further aspect of the invention, the images are divided into blocks and the indication of image content is computed for each block. More specifically the indication of image content is computed from a histogram for each block. Thereafter, the step of evaluating the indication of image content and  
20 the time of capture comprises comparing one or more blocks of one image, using a histogram intersection metric, to corresponding blocks of another image and using the time difference between capture of the two images to determine whether the images are duplicate images. Moreover, the step of computing an indication of image content may include dividing each image into blocks, wherein one or more  
25 blocks represent a foreground area of the images, and computing an indication of image content in each block and in the foreground areas of each block.

Consequently, according to this method image content is analyzed to determine duplicates by the similarity of content of the images and the time of exposure. If the image contents are similar and the time difference between  
30 exposures is within a certain threshold, then the images are duplicates. If no time and date information is available, the image content alone can be used.

These and other aspects, objects, features and advantages of the present invention will be more clearly understood and appreciated from a review

of the following detailed description of the preferred embodiments and appended claims, and by reference to the accompanying drawings.

### **ADVANTAGEOUS EFFECT OF THE INVENTION**

5           When two pictures are identified as duplicates, one of them is removed and will not appear in the album. Since the consumer will not be satisfied if the automatic albuming process removes a picture that should have been in the album, the precision of the duplicate detection algorithm has to be high. The present invention provides for an automatic albuming process with  
10   such a high precision rate.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

FIG. 1 is a block diagram illustrating an overview of the present invention;

15           FIG. 2 shows a comparison of Recall versus Precision curves determined for different block sizes, including 2x2, 3x3 and 4x4 block sizes;

FIG. 3 shows the results of the comparisons between a 3x3 blocks technique with and without the use of time information;

20           FIG. 4 shows the area covered by a 3x3 arrangement of blocks, including foreground areas represented by blocks 5 and 8;

FIGS. 5A and B show examples of color histograms for the central area covered by block 5 as shown in Figure 4;

25           FIG. 6 shows the comparisons between various 3x3 blocks techniques showing the effect of using (and not using) time information, using the average of the nine histogram intersection values, and using the rules shown with regard to Figure 1; and

FIG. 7 shows a table summarizing the improvements made according to the invention to the duplicate detection method.

30           FIG. 8 is a block diagram of a computer system for implementing the present invention.

## DETAILED DESCRIPTION OF THE INVENTION

In the following description, the present invention will be described in the preferred embodiment as a software program. Those skilled in the art will readily recognize that the equivalent of such software may also be constructed in hardware. Given the system and method as shown and described according to the invention in the following materials, software not specifically shown, described or suggested herein that is useful for implementation of the invention is conventional and within the ordinary skill in such arts. Still further, as used herein, computer readable storage medium may comprise, for example; magnetic storage media such as a magnetic disk (such as a floppy disk or a hard drive) or magnetic tape; optical storage media such as an optical disc, optical tape, or machine readable bar code; a solid state electronic storage device such as random access memory (RAM), or read only memory (ROM); or any other physical device or medium employed to store a computer program.

Referring now to Fig. 1, there is illustrated a flow diagram illustrating an overview of the present invention. Digitized images are input into a computer system in step S10, where a software program will classify them into distinct categories. For instance, the images will be ranked in chronological order by analyzing the time of capture of each image (date may also be used to isolate the time by day, and in the absence of time, date alone can provide a gross estimate of chronological order). The time of capture of each picture may be extracted, for example, from the encoded information on the film strip of the Advanced Photo System (APS) images, or from information available from some digital cameras. Furthermore, each image is divided into NxN blocks (with N = 2, 3 or 4 in typical implementations). In the preferred embodiment, the image is divided into 3x3 blocks, as shown in Figure 4. For each block, an indication of image content is computed; more specifically, a color histogram is computed for each block (exemplary histograms of the center block are shown for two similar images in Figures 5A and 5B). Then each block of one image is compared, using a conventional histogram intersection metric, to the corresponding block of another image. (The histogram intersection metric is described in the aforementioned Serial Number 09/163,618, which is incorporated herein by reference.)

At the beginning of the process, two flags IsDup and ForegroundDup are set to true and false, respectively, in a step S11. Then, in a step S12, the individual histogram intersection value for each block (except for the center block) is checked to determine whether it is higher than a certain threshold  $T_1$ , and the number of intersection values below the threshold  $T_1$  are counted. If the number of blocks with intersection values below the threshold  $T_1$  are found in a step S14 to be greater than a certain number  $N$ , then the flag IsDup is set to false. Next, the foreground average histogram intersection value is computed in step S15 for a foreground area of the image (where foreground, for purposes of a 3x3 block, is defined as blocks 5 and 8, as shown in Figure 4). If the foreground average histogram intersection value is found in a step S16 to be less than a certain threshold  $T_2$ , then the flag IsDup is set to false. If the foreground average histogram intersection value is found in a step S18 to be higher than a certain threshold  $T_3$ , then the flag ForegroundDup is set to true.

Next, the average of all histogram intersection values for all blocks of each image calculated and checked in a step S20 to determine whether that average histogram intersection value is greater than a threshold  $T_4$  (with the flag IsDup being true), and further checked in a step S22 to determine whether that average histogram intersection value is greater than another threshold  $T_5$  (where  $T_5 < T_4$ , and where the flag ForegroundDup is true). If neither one of these conditions is satisfied, then the decision is made that the image(s) are not duplicate images. If either one of these conditions is satisfied, the decision regarding duplicates is not made at this point. Then, in step S24 the time difference between the two images is checked to determine whether it is less than a certain threshold  $T_6$ , and further in step S28 checked to determine whether it is less than another threshold  $T_8$  (where  $T_6 < T_8$ ). If the condition in step 24 is satisfied, or if the condition in step 28 is satisfied in combination with the aforementioned average of all histogram intersection values being found greater in a step 26 than another threshold  $T_7$  (with  $T_5 < T_4 < T_7$ ), then the decision is made that the image(s) are duplicate images. Otherwise, the images are determined not to be duplicate images.

In summary, the rules for determining whether an image is a duplicate of another image is as follows:

- a) If S14 and S16 are not satisfied, and S20 and S24 are satisfied, the pictures are duplicates.
- b) If S14 and S16 are not satisfied, and S20, S26 and S28 are satisfied, the pictures are duplicates.
- 5 c) If S18, S22, and S24 are satisfied the pictures are duplicates.
- d) If S18, S22, S26 and S28 are satisfied the pictures are duplicates.
- e) Otherwise the pictures are not duplicates.

10

The goal of the duplicate detection algorithm outlined in Figure 1 is to identify whether two pictures are so similar that a consumer would only put one of them in an album. A workable definition of a duplicate is as follows: duplicates are defined as two photographs that have the same content and  
15 composition, as well as being taken from the same angle and range. Typically, substantially exactly the same content and composition are required to establish a duplicate, although a small variation in angle and/or range may be tolerated. Consequently, the duplicate detection method is trying to recall the "almost identical" pictures, i.e., the two pictures should have almost the same colored  
20 pixels at almost the same locations. Obviously, a pixel to pixel comparison will not work well, since there will always exist a small shift; on the other hand, a global color histogram approach will not be accurate enough due to the lack of information about the pixels' locations. The block histogram approach gives semi-local information on the pixel colors and some information of location  
25 within the image. The blocks cannot be too big so that the small shift between the images will not affect the accuracy. In the experiment as described below, we decided to divide the image into 3x3 blocks, as shown in Figure 4.

### EXPERIMENT

30 We have found that date and time information can be very useful in achieving the goal of the duplicate detection algorithm. Indeed, according to the aforementioned definition of duplicates, it would clearly appear that such pictures are usually taken within a short period of time. By setting different

thresholds in coordination with the image similarity values, as shown and described in connection with Figure 1, we were able to improve the quality of the duplicate detection significantly.

In order to benchmark and verify the algorithm, a third party ground truth database was set up. Four hundred forty three (443) pictures were carefully chosen from the database. The pictures are all duplicates candidates, including a lot of pictures that are obviously not duplicates for a human eye but might be for a machine readable apparatus (same picture but different people, etc.). The database contains about 270 pairs. The third party ground truth has been based on the participation of ten observers. Each observer was given a definition of what are duplicates pictures plus some explanation of how to make a decision. In addition, it was explained to the observers not to put themselves in the place of the consumer, that is, not to develop a like or dislike for the images, but to put themselves in the place of a third party person. Ideally, the interest of the pictures to the observers should not influence their decision whether the images are duplicates. The observers were told that their input will be used for the benchmarking of a duplicate detection system.

The output of the duplicate detection method is binary, meaning that a picture is flagged as either a duplicate or not a duplicate. Nevertheless, what the ground truth study provides, for each pair of pictures, is a probability of the pair being duplicates. The metric used to benchmark the method is based on a Recall variable versus a Precision variable, where:

$$Recall = \frac{\#correct}{\#correct + \#missed} \text{ and } Precision = \frac{\#correct}{\#correct + \#false\ positive} \text{ (Eq.1)}$$

The first step was to determine the number of blocks into which the images would be divided. In the aforementioned Serial Number 09/163,618, a block-based histogram technique was used for event classification but it involved a much larger number of blocks. In the comparative illustration of Figure 2, Recall versus Precision curves were generated for the 2x2 blocks, 3x3 blocks, 4x4 blocks techniques and for a block-based histogram technique involving a larger number of blocks. For the several NxN blocks techniques, an average of the

histogram intersections is computed and a threshold is applied. The Recall versus Precision curves are obtained by varying the threshold for each technique. The results of the comparison are shown in Figure 2. Figure 2 shows that the new approach with smaller blocks is better than a block-based histogram technique involving a larger number of blocks for the detection of duplicates. It also appears clear that the results of the 3x3 blocks and the 4x4 blocks approaches exceed the results of the 2x2 blocks approach. We decided to use the 3x3 blocks approach for the following reasons: 1) the results are slightly better than with the use of 4x4 blocks; and 2) the 3x3 blocks approach has the advantage of having a middle block, which for the majority of the pictures is likely to contain the main subject.

The next step was to determine the influence of the date and time information on the quality of the results. Date and time information turned out to be very relevant information, and enabled us to improve the precision of the results significantly, even though only 57% of the pictures had date and time information. Adaptive thresholds were set up for the time, all optimized for this database.

Figure 3 shows the results of a comparison between a 3x3 blocks technique without the use of time information and a 3x3 blocks technique with the use of time information. Figure 3 shows that for reasonable Recall, between 0.65 and 0.75, the precision obtained with date and time included is much higher with an average improvement of 0.065, or 9%. Finally, we optimized the thresholding technique on each block. In addition to the threshold on the average histogram intersections, a threshold on each histogram intersection and a threshold on the average histogram intersections of blocks 5 and 8 have been set up. The blocks 5 and 8 are very likely to contain the main subject. They generally represent the foreground of the image. Figure 4 shows the area covered by blocks 5 and 8.

Figures 5A and 5B illustrate how the blocks of each image are compared to each other using color histograms, and more specifically show how the RGB color histograms of the center block of each picture are compared. Even though there is a slight shift between the two images, these figures show that the color histograms for the center block are basically the same. The only noticeable difference is the observable peak in the right part of the histogram of the block of

picture B (Figure 5B). This peak corresponds to the small amount of water present in the center block of picture B that is not present in the center block of picture A (Figure 5A) due to the shift. This difference will not be a significant influence on the overall intersection value between these histograms.

5                   Figure 6 shows comparisons between the use of several 3x3 block histogram techniques, one without the use of time information and the others either with the use of time information in combination with the average of the nine histogram intersection values or with time information in combination with the set of rules (or a subset) described above in relation to Figure 1. More  
10 specifically, the curve in Figure 6 labeled *3x3+time.3*, which follows the process shown in Figure 1, shows the improvements realized by the application of the set of rules described above in relation to Figure 1. The best solution was obtained using the following set of thresholds:  $T_1=0.54$ ;  $T_2=0.43$ ;  $T_3=0.57$ ;  $T_4=0.495$ ;  $T_5=0.62$ ;  $T_6=2$ ;  $T_7=0.62$ ;  $T_8=8$ ;  $N=2$ . The result achieved is Recall = 0.68 and  
15 Precision = 0.81. This represents an improvement of 0.045 or 6% in Precision at equal Recall compared with the basic technique using time information.

Figure 7 shows a table summarizing the improvements made to the duplicate detection method, in particular showing that the technique according to the invention achieves a significant improvement compared with the block-based  
20 histogram technique (an increase of 5.4% in Recall and 23.8% in Precision can be observed).

-----

In another embodiment, If no time and date information is available, the block histogram analysis alone provides a method for determining  
25 duplicates in an automatic albing system. Referring to a subset of steps in Figure 1, after inputting the images in step S10 determine in step S12 if each individual histogram intersection value (except the center block) is higher than a certain threshold  $T_1$ , and count the number of intersection values below that threshold. Check whether the number of intersection values below that threshold  
30 are found in step S14 to be greater than a certain number N. Compute in step S15 the average histogram intersection of blocks 5 and 8. Check in step S16 whether that average is less than a certain threshold  $T_2$ , and check in step S18 whether that average is higher than a certain threshold  $T_3$ . Compute the average of the all

histogram intersection values and check in step S20 whether the average of all histogram intersection values is higher than a threshold  $T_4$ . Finally, check in step S22 whether the average of all histogram intersection values is higher than a threshold  $T_5$  (with  $T_5 < T_4$ ). Then, these various computations are applied as  
5 shown in the appropriate parts of Figure 1 to determine if the input images are duplicates.

While the overall methodology of the invention is described above, the invention can be embodied in any number of different types of systems and executed in any number of different ways, as would be known by one ordinarily  
10 skilled in the art. It facilitates understanding to note that the present invention is preferably utilized on any well-known computer system, such as a personal computer. It is also instructive to note that the images may be either directly input into the computer system (for example by a digital camera) or digitized before input into the computer system (for example by scanning). For example, as  
15 illustrated in Figure 8, a typical hardware configuration of an information handling/computer system useful in implementing the invention preferably has at least one processor or central processing unit (CPU) 100. The CPU 100 is interconnected via a system bus 101 to a random access memory (RAM) 102, a read-only memory (ROM) 103, an input/output (I/O) adapter 104 (for connecting  
20 peripheral devices such as disk units 105 and tape drives 106 to the bus 101), a communication adapter 107 (for connecting an information handling system to a data processing network, such as the Internet), a user interface adapter 108 (for connecting peripherals 109, 110, 111 such as a keyboard, mouse, digital image input unit (e.g., a scanner or a camera), microphone speaker and/or other user  
25 interface device to the bus 101), a printer 112 and a display adapter 113 (for connecting the bus 101 to a display device 114). The invention could be implemented using the structure shown in Figure 8 by including the inventive method within a computer program stored, e.g., on the storage device 105. Such a computer program would act on a time series of image frames supplied through  
30 the interface adapter 108 or through the network connection 107 in order to detect duplicates. The system would then automatically produce the desired digital image frame output (without duplicates) on the display 114, the printer 112 or sent back to the network 107.

The invention has been described with reference to a preferred embodiment. However, it will be appreciated that variations and modifications can be effected by a person of ordinary skill in the art without departing from the scope of the invention.

## PARTS LIST

100	CPU
101	bus
102	RAM
103	ROM
104	I/O adapter
105	disk unit
106	tape drive
107	communication adapter
108	interface adapter
109	keyboard
110	mouse
111	digital image input unit
112	printer
113	display adapter
114	display device

**WHAT IS CLAIMED IS:**

1. A method for detecting duplicate images comprising the steps of:  
  
providing at least two images captured at determinable times;  
computing an indication of image content for each image;  
determining the time of capture of each of the images; and  
evaluating the indication of image content and the time of capture to determine whether the images are duplicate images.
2. The method as claimed in claim 1 wherein the step of computing an indication of image content comprises:  
  
dividing each image into blocks; and  
computing an indication of image content in each block.
3. The method as claimed in claim 2 wherein each image is divided into 4x4 or fewer blocks.
4. The method as claimed in claim 3 wherein each image is divided into 3x3 blocks.
5. The method as claimed in claim 2 wherein the step of computing an indication of image content in each block comprises computing a histogram for each block.
6. The method as claimed in claim 5 wherein the step of evaluating the indication of image content and the time of capture comprises comparing one or more blocks of one image, using a histogram intersection metric, to corresponding blocks of another image and using the time difference between capture of the two images to determine whether the images are duplicate images.

7. The method as claimed in claim 1 wherein the step of computing an indication of image content comprises:

dividing each image into blocks, wherein one or more blocks represent a foreground area of the images; and  
computing an indication of image content in each block and in the foreground areas of each image.

8. A method for detecting duplicate images comprising the steps of:

- (a) providing a plurality of images captured at determinable times;
- (b) dividing each image into an X number of blocks, wherein one or more blocks represent a central area and a foreground area;
- (c) computing histograms for each block of each image, and block histogram intersection values obtained from comparisons between histograms from corresponding blocks from each image;
- (d) determining whether each block histogram intersection value for at least those blocks surrounding the central area is higher than a threshold  $T_1$ , and determining whether the number of intersection values below the threshold  $T_1$  are not greater than a certain number N;
- (e) computing an average histogram intersection value of the foreground area, and determining whether the average block histogram intersection value of the foreground area is not lower than a threshold  $T_2$ ;
- (f) determining whether the average histogram intersection value of the foreground is higher than a threshold  $T_3$ ;
- (g) determining whether an average of the X number of block histogram intersection values is higher than a threshold  $T_4$ ;
- (h) determining whether the average of the X number of block histogram intersection values is higher than a threshold  $T_5$ ;
- (i) determining whether the time difference between capture of the images is less than a threshold  $T_6$ ;
- (j) determining whether the average of the X number of block histogram intersection values is higher than a threshold  $T_7$ ; and

- (k) determining whether the time difference between the capture of the images is less than a threshold  $T_8$ ; and
- (l) utilizing the determinations made in steps (d) through (k) to determine if any of the images are duplicates.

9. The method as recited in claim 8 wherein said step (h) further provides that  $T_5 < T_4$ .

10. The method as recited in claim 8 wherein said step (j) further provides that  $T_5 < T_4 < T_7$ .

11. The method as recited in claim 8 wherein said step (k) further provides that  $T_6 < T_8$ .

12. The method as recited in claim 8 wherein said step (b) comprises dividing each image into a configuration of 4x4 or fewer blocks.

13. The method as recited in claim 12 wherein said step (b) comprises dividing each image into a configuration of 3x3 blocks.

14. A method for detecting duplicate images comprising the steps of:

- (a) providing a plurality of images;
- (b) dividing each image into an X number of blocks, wherein one or more blocks represent a central area and a foreground area;
- (c) computing histograms for each block, and block histogram intersection values obtained from comparisons between histograms from corresponding blocks from each image;
- (d) determining whether each block histogram intersection value for at least those blocks surrounding the center block is higher than a threshold  $T_1$ , and determining whether the number of intersection values below the threshold  $T_1$  are not greater than a certain number N;

- (e) computing an average histogram intersection value of the foreground area, and determining whether the average block histogram intersection value of the foreground area is not lower than a threshold  $T_2$ ;
- (f) determining whether the average histogram intersection value of the foreground is higher than a threshold  $T_3$ ;
- (g) determining whether an average of the X number of block histogram intersection values is higher than a threshold  $T_4$ ;
- (h) determining whether the average of the X number of block histogram intersection values is higher than a threshold  $T_5$ ; and
- (i) utilizing the determinations made in steps (d) through (h) to determine if any of the images are duplicates.

15. The method as recited in claim 14 wherein said step (h) further provides that  $T_5 < T_4$ .

16. The method as recited in claim 14 wherein said step (b) comprises dividing each image into a configuration of 4x4 or fewer blocks.

17. The method as recited in claim 16 wherein said step (b) comprises dividing each image into a configuration of 3x3 blocks.

18. A computer program product for detecting duplicate images comprising: a computer readable storage medium having a computer program stored thereon for performing the steps of:

- providing at least two images captured at determinable times;
- computing an indication of image content for each image;
- determining the time of capture of each of the images; and
- evaluating the indication of image content and the time of capture to determine whether the images are duplicate images.

19. The computer program product as claimed in claim 18 wherein the step of computing an indication of image content comprises:  
dividing each image into blocks; and

computing an indication of image content in each block.

20. The computer program product as claimed in claim 19 wherein each image is divided into 4x4 or fewer blocks.

21. The computer program product as claimed in claim 20 wherein each image is divided into 3x3 blocks.

22. The computer program product as claimed in claim 19 wherein the step of computing an indication of image content in each block comprises computing a histogram for each block.

23. The computer program product as claimed in claim 22 wherein the step of evaluating the indication of image content and the time of capture comprises comparing one or more blocks of one image, using a histogram intersection metric, to corresponding blocks of another image and using the time difference between capture of the two images to determine whether the images are duplicate images.

24. The computer program product as claimed in claim 18 wherein the step of computing an indication of image content comprises:  
dividing each image into blocks, wherein one or more blocks represent a foreground area of the images; and  
computing an indication of image content in each block and in the foreground areas of each image.

### ABSTRACT

Duplicate images are detected by providing at least two images captured at determinable times; computing an indication of the image content for each image; determining the time of capture of each of the images; and evaluating

5 the indication of image content and the time of capture to determine whether the images are duplicate images. The images are divided into blocks and the indication of image content is computed from a histogram for each block. Thereafter, the step of evaluating the indication of image content and the time of capture comprises comparing one or more blocks of one image, using a histogram

10 intersection metric, to corresponding blocks of another image and using the time difference between capture of the two images to determine whether the images are duplicate images.

US5N 09/163,618  
filed 9/30/98

Original

Application Based on

Docket **78311PCW**

Inventors: Alexander C. Loui and Eric S. Pavie

**A METHOD FOR AUTOMATICALLY CLASSIFYING IMAGES  
INTO EVENTS**

Assistant Commissioner for Patents,  
ATTN: BOX PATENT APPLICATION  
Washington, D. C. 20231

Express Mail Label No.: EM364958645US

Date: Sept. 30, 1998

# **A METHOD FOR AUTOMATICALLY CLASSIFYING IMAGES INTO EVENTS**

## **FIELD OF THE INVENTION**

5           The invention relates generally to the field of image processing having image understanding that automatically classifies pictures by events and the like and, more particularly, to such automatic classification of pictures by time and date analysis and by block-based analysis which selectively compares blocks of the images with each other.

10

## **BACKGROUND OF THE INVENTION**

Pictorial images are often classified by the particular event, subject or the like for convenience of retrieving, reviewing, and albing of the images. Typically, this has been achieved by manually segmenting the images, or by the  
15 below-described automated method. The automated method includes grouping by color, shape or texture of the images for partitioning the images into groups of similar image characteristics.

Although the presently known and utilized methods for partitioning images are satisfactory, there are drawbacks. The manual classification is  
20 obviously time consuming, and the automated process, although theoretically classifying the images into events, is susceptible to miss-classification due to the inherent inaccuracies involved with classification by color, shape or texture.

Consequently, a need exists for overcoming the above-described drawbacks.

25

## **SUMMARY OF THE INVENTION**

The present invention is directed to overcoming one or more of the problems set forth above. Briefly summarized, according to one aspect of the present invention, the invention resides in a method for automatically classifying  
30 images into events, the method comprising the steps of: receiving a plurality of images having either or both date and/or time of image capture; determining one

or more largest time differences of the plurality of images based on clustering of the images; and separating the plurality of images into the events based on having one or more boundaries between events which one or more boundaries correspond to the one or more largest time differences.

5                   These and other aspects, objects, features and advantages of the present invention will be more clearly understood and appreciated from a review of the following detailed description of the preferred embodiments and appended claims, and by reference to the accompanying drawings.

10                   **ADVANTAGEOUS EFFECT OF THE INVENTION**

                  The present invention has the advantage of improved classification of images by utilizing both date and time information and block-based comparison that checks for similarity of subject and background in the images. If date and time information is not available, then the block-based analysis may be used as the  
15                   sole basis for classification.

**BRIEF DESCRIPTION OF THE DRAWINGS**

                  Fig. 1 is a block diagram illustrating an overview of the present invention;

20                   Fig. 2 is a block diagram illustrating a date and time clustering technique of the present invention;

                  Fig. 3 is a graph illustrating a scaling function used to map the result of the 2-means clustering;

25                   Fig. 4 is a graph illustrating a typical result of the scaling function of Fig. 3;

                  Fig. 5 is a diagram illustrating a block diagram of an event boundary checking after the date and time clustering;

                  Fig. 6 is a diagram illustrating grouping of images within each event based on content;

30                   Fig. 7 is a block diagram of a group-merging step of the present invention;

Fig. 8 is a block diagram of image re-arrangement within each group;

Fig. 9 is a flowchart of block-based histogram correlation technique;

5 Fig. 10 is diagram illustrating the comparison between block histogram;

Fig. 11 is diagram of an example of best intersection mapping for three segment analysis; and,

10 Fig. 12 is an illustration of shift detection within the block based histogram correlation.

### **DETAILED DESCRIPTION OF THE INVENTION**

In the following description, the present invention will be described in the preferred embodiment as a software program. Those skilled in the art will  
15 readily recognize that the equivalent of such software may also be constructed in hardware.

Still further, as used herein, computer readable storage medium may comprise, for example; magnetic storage media such as a magnetic disk (such as a floppy disk) or magnetic tape; optical storage media such as an optical disc,  
20 optical tape, or machine readable bar code; solid state electronic storage devices such as random access memory (RAM), or read only memory (ROM); or any other physical device or medium employed to store a computer program.

In addition, the term "event" is defined herein as a significant occurrence or happening as perceived by the subjective intent of the user of the  
25 image capture device.

Before describing the present invention, it facilitates understanding to note that the present invention is preferably utilized on any well-known computer system, such a personal computer. Consequently, the computer system will not be discussed in detail herein. It is also instructive to note that the images  
30 are either directly input into the computer system (for example by a digital

camera) or digitized before input into the computer system (for example by scanning).

Referring to now Fig. 1, there is illustrated a flow diagram illustrating an overview of the present invention. Digitized images are input into the computer system where a software program of the present invention will  
5     classify them into distinct categories. The images will first be ranked S10 in chronological order by analyzing the time and date of capture of each image. The date and/or time of capture of each picture may be extracted, for example, from the encoded information on the film strip of the Advanced Photo System (APS)  
10    images, or from information available from some digital cameras. The representations of the images will then be placed S20 into one of a plurality of distinct events by a date and time clustering analysis that is described below. Within each event, the contents of the images are analyzed S20 for determining whether images closest in time to an adjacent event should be maintained in the  
15    event as defined by the clustering analysis, or the adjacent events merged together. After the images are defined into events, a further sub-classification (grouping) within each event is performed. In this regard, the images within each event will then be analyzed by content S30 for grouping images of similar content together, and then the date and time S30 for further refining the grouping.

20             The event segmentation S20 using the date and time information is by a k-means clustering technique, as will be described in detail below, which groups the images into events or segments. A boundary check is then performed on the segments S20 for verifying that the boundary images should actually be grouped into the segment identified by the clustering, as will also be described  
25    below.

These groups of images are then sent to a block-based histogram correlator S30 for analyzing the content. For each event or segment sent to the correlator, a content or subject grouping S30 is performed thereon for further sub-classifying the images by subject within the particular event segment. For  
30    example, within one event, several different subjects may appear, and these subject groupings define these particular subjects. The subject grouping is based

primarily on image content, which is performed by a block-based histogram correlation technique. This correlation compares portions of two images with each other, as will also be described in detail below. The result of the ranking is the classification of images of each segment into distinct subject groupings. The  
5 date and time of all the images within each subject grouping are then compared to check whether any two or more subject grouping can be merged into a single subject grouping S30.

A refinement and subject re-arrangement analysis S40 will further improve the overall classification and the subject grouping by rearranging certain  
10 images within a subject group.

Referring to Fig. 2, there is shown an exploded block diagram illustrating the data and time clustering technique S20. First, the time interval between adjacent pictures (time difference) is computed S20a. A histogram of the time differences is then computed S20b, an example of which is shown in block  
15 10. The abscissa of the histogram is preferably the time in minutes, and the ordinate of the histogram is the number of pictures having the specified time difference. The histogram is then mapped S20c to a scaled histogram using a time difference scaling function, which is shown in Fig. 3. This mapping is to take the pictures with small time difference and substantially maintain its time difference,  
20 and to take pictures with a large time difference and compress their time difference.

A 2-means clustering is then performed S20d on the mapped time-difference histogram for separating the mapped histogram 10 into two clusters based on the time difference; the dashed line represents the separation point for the  
25 two clusters. For further details of 2-means clustering, Introduction to Statistical Pattern Recognition, 2<sup>nd</sup> edition by Keinosuke Fukunaga 1990 may be consulted, and therefore, the process of 2-means clustering will not be discussed in detail herein. Referring briefly to Fig. 4, the result of 2-means clustering is the segmentation of the histogram into two portions 10a and 10b. Normally, events  
30 are separated by large time differences. The 2-means clustering, therefore, is to define where these large time differences actually exist. In this regard, the right

portion 10b of the 2-means clustering output defines the large time differences that correspond to the event boundaries.

Referring to Fig. 5, there is illustrated an example of boundary checking between events. For two consecutive events  $i$  and  $j$ , a plurality of block-based, histogram comparisons are made to check if the pictures at the border of one event are different from the pictures at the border of the other event. If the comparison of content is similar, the two segments are merged into one segment. Otherwise, the segments are not merged. Preferably, the comparisons are performed on the three border images of each event ( $i_3, i_4, i_5$  with  $j_1, j_2, j_3$ ), as illustrated in the drawing. For example, image  $i_5$  is compared with image  $j_1$  and etc. This block-based histogram comparison technique will be described in detail hereinbelow.

Referring to Fig. 6, there is illustrated an overview of subject (content) grouping for each segmented event. Within each segmented event  $i$ , adjacent pictures are compared (as illustrated by the arrows) with each other using the below-described, block-based histogram technique. For example, the block-based histogram technique may produce five subject groupings (for example groups 1-5) from the one event  $i$ , as illustrated in the drawing. The arrangement of the subject grouping is stored for future retrieval during the subject arrangement step s40. In particular, the subject grouping having a single image is stored (for example groups 2, 3, and 5).

Referring to Fig. 7, after the grouping by content, a time and date ordering is performed on the groupings for merging groups together based on a time and date analysis. A histogram of the time difference between adjacent images in the event is computed, similar to Fig. 4. A predetermined number of the largest time differences (for example boundary  $a_{12}$ ) are compared with the boundaries (for example boundaries  $b_{12}, b_{23}, b_{34}, b_{45}$ ) of the subject grouping determined by the block-based analysis. The predetermined number of largest time differences are determined by dividing the total number of images within an event by the average number of picture per group (four is used in the present invention). If the boundary of the subject grouping matches the boundary based

on the chosen time differences, the subject groupings will not be merged. If there is not a match between the two boundaries, the subject groupings having a boundary not having a matched time difference in the histogram will be merged into a single subject grouping (for example groups  $b_1, b_2, b_3$  merged into resulting group  $c_1$ ).

Referring to Fig. 8, there is illustrated a diagram of image re-arrangement within each group. The arrangement of the initial subject groupings is retrieved for identifying subject groupings that contain single images (for example the groups with a single image of Fig. 6 –groups 2, 3, and 5 that are re-illustrated as groups 2, 3, and 5 in Fig. 8). Any single images from the same subject grouping that are merged as identified by the merged subject grouping (for example, groups  $c_1$  and  $c_2$  of Fig. 7) are compared with all other images in the merged subject grouping, as illustrated by the arrows. This comparison is based on block-based histogram analysis. If the comparisons are similar, these images will be re-arranged so that the similar images are located adjacent each other, for example groups  $d_1$  and  $d_2$ .

Further refinement is done by comparing any group that still contains a single image after the above procedure, with all the images in the event. This is to check if these single image groups can be better arranged within the event grouping. This comparison is similar to the subject re-arrangement step of Fig. 8.

Referring to Fig. 9, there is illustrated a flowchart of the block-based histogram correlation used in the above analyses. First, a histogram of the entire image of both images is computed S50, a global histogram. A comparison of the two histograms is performed by histogram intersection value S60 illustrated the following equation:

$$Inter(R, C) = \frac{\sum_{i=1}^n \min(R_i, C_i)}{\sum_{i=1}^n R_i}$$

1

where R is the histogram of the reference image, C is the histogram of the candidate image, and n is the number of bins in the histogram. If the intersection is under a threshold S65, preferably 0.34, although other thresholds may be used, the images are different. If the threshold is met or exceeded S65, then a block-based histogram correlation will be performed S70. In this regard, each image will be divided into blocks of a given size, preferably 32 x 32 pixels in the present invention. It is instructive to note that those skilled in the art may vary the block size depending on the resolution of the image without departing from the scope of the invention. For each block, a color histogram is computed. Referring to Fig. 10, if one image is considered a reference image and one image a candidate image, the images are compared in the following way. Each block 20 of the reference image is compared to the corresponding block 30 of the candidate image and to the adjacent blocks 40, 8 blocks in the present invention.

Referring to Fig. 9, the block histograms between the reference image and the candidate image are compared using the histogram intersection equation defined above S80. The average intersection value is derived by computing the average of the best intersection values from each of the block comparisons S90. This average intersection value will be compared to a low threshold (preferably 0.355), and a high threshold (preferably 0.557). If the average intersection value is below the low threshold S95, the two images are considered different. If the average intersection value is above the high threshold S96, then the two images are considered similar. If the average intersection value is between these two thresholds, further analysis will be performed as described below (3-segment average intersection map S100).

Referring to both Figs. 9 and 11, a 3-segment analysis will be performed to determine if the two images may contain a similar subject. This is performed by first forming a map 60 which contains the average of the two highest intersection values of each of the block comparisons; for example, 9 comparisons were performed in the illustration of Fig. 10, the average of the highest two will be used for map 60. Fig. 11 illustrates, for example, a 9 x 6 block although it should be understood that the map size depends on the size of the image. This map is

divided into three parts: the left portion 70a, the center portion 70b, and the right portion 70c. If the average intersection value of the center portion 70b is higher than a threshold (preferably 0.38) S105, the two images may contain a very similar subject in the center portion 70b of the image, and the two images may be

- 5 considered to be similar by subject. In addition, the comparisons of the histogram will be performed with the reference and candidate images reversed. If the two images are similar both methods should give substantially similar correlation; obviously if they are different, the results will not be similar. The images are then checked S110 to determine if there is a high intersection value in one of the
- 10 directions, right, left, up, and down.

- Referring to Figs. 9 and 12, shift detection is used to determine the case when the two images 90 and 100 (of two different sizes in the drawing) have very similar subject that appears in different locations of the image. For example, the main subject may be situated in the center of one image and to the left-hand
- 15 side of the other image. Such a shift can be determined by recording both the best intersection values of the reference blocks, as well as the coordinates of the corresponding candidate blocks. This is achieved by comparing the intersection values of the blocks in four directions (right, left, up, and down). The entire image will be shifted by one block (as illustrated by the arrows) in one of the directions
- 20 (right in the drawing) where the best intersection value is the highest. The above analysis and the shift can be repeated S120 to check for similarity.

- The invention has been described with reference to a preferred embodiment. However, it will be appreciated that variations and modifications can be effected by a person of ordinary skill in the art without departing from the
- 25 scope of the invention.

**CLAIMS:**

1. A method for automatically classifying images into events, the method comprising the steps of:
  - (a) receiving a plurality of images having either or both date and/or time of image capture;
  - (b) determining one or more largest time differences of the plurality of images based on clustering of the images; and,
  - (c) separating the plurality of images into the events based on having one or more boundaries between events which one or more boundaries correspond to the one or more largest time differences.
2. The method as in claim 1, wherein step (b) includes computing a time difference histogram and performing a 2-means clustering on the time difference histogram for defining the one or more boundaries.
3. The method as in claim 2, wherein step (b) further includes mapping the time difference histogram through a time difference scaling function before performing the 2-means clustering.
4. The method as in claim 2, wherein step (c) includes checking the images adjacent the one or more boundaries for similarity by comparing content of the images.
5. The method as in claim 4, wherein step (c) includes checking the images adjacent the one or more boundaries for similarity by using a block-based histogram correlation technique.
6. The method as in claim 5 further comprising step (d) dividing the events into subject grouping by using an image content analysis.

7. The method as in claim 6, wherein step (d) includes dividing the events into subject grouping by using a block-based histogram technique.

8. A method for automatically classifying images into events, the method comprising the steps of:

- (a) receiving a plurality of images arranged in chronological order;
- (b) dividing the images into a plurality of blocks; and,
- (c) grouping the images into subject grouping based on block-based histogram correlation which includes computing a color histogram of each block and computing a histogram intersection value which determines the similarity between blocks.

9. The method as in claim 8, wherein step (c) includes comparisons of two of the images by shifting one of the images in a desired direction based on the intersection value and then computing the block based correlation.

10. The method as in claim 9, wherein step (c) includes forming a map that contains two best intersection values of each of the block comparisons ; dividing the map into three portions; and then comparing center portions for similarity.

ABSTRACT

A method for automatically classifying images into events, the method includes the steps of: receiving a plurality of images having either or both  
5 date and/or time of image capture; determining one or more largest time differences of the plurality of images based on clustering of the images; and separating the plurality of images into the events based on having one or more boundaries between events which one or more boundaries correspond to the one or more largest time differences.

10

USSN 09/197,365  
filed 11/20/98

ORIGINAL

Application Based on

Docket 78713PCW

Inventors: Alexander C. Loui and Eric S. Pavie

**A METHOD FOR AUTOMATICALLY COMPARING CONTENT OF  
IMAGES FOR CLASSIFICATION INTO EVENTS**

Assistant Commissioner for Patents,  
ATTN: BOX PATENT APPLICATION  
Washington, D. C. 20231

Express Mail Label No.: EM364958685US

Date: November 20, 1998

**A METHOD FOR AUTOMATICALLY COMPARING  
CONTENT OF IMAGES FOR CLASSIFICATION INTO  
EVENTS**

**5                    CROSS-REFERENCE TO RELATED APPLICATIONS**

The present application is related to U.S. Application Serial Number 09/163,618, filed September 30, 1998, by Alexander C. Loui and Eric S. Pavie, and entitled, "A METHOD FOR AUTOMATICALLY CLASSIFYING IMAGES INTO EVENTS".

10

**FIELD OF THE INVENTION**

The invention relates generally to the field of image processing having image understanding that automatically classifies pictures by events and the like and, more particularly, to such automatic classification of pictures by block-based analysis which selectively compares blocks of the images with each other.

15

**BACKGROUND OF THE INVENTION**

Pictorial images are often classified by the particular event, subject or the like for convenience of retrieving, reviewing, and albing of the images. Typically, this has been achieved by manually segmenting the images, or by the below-described automated method. The automated method includes grouping by color, shape or texture of the images for partitioning the images into groups of similar image characteristics.

20

Although the presently known and utilized methods for partitioning images are satisfactory, there are drawbacks. The manual classification is obviously time consuming, and the automated process, although theoretically classifying the images into events, is susceptible to miss-classification due to the inherent inaccuracies involved with classification by color, shape or texture.

25

Consequently, a need exists for overcoming the above-described drawbacks.

30

### **SUMMARY OF THE INVENTION**

The present invention is directed to overcoming one or more of the problems set forth above. Briefly summarized, according to one aspect of the present invention, the invention resides in a method for comparing image content of first and second images, the method comprising the steps of: (a) extracting a portion of both the first and second images both of which portions are determined to include a main subject area of each image; (b) dividing the main subject area of the images into a plurality of blocks; (c) computing a color histogram for one block in each image; (d) computing a histogram intersection value between the block of the first image and the block of the second image; and (e) determining a first threshold value for the computed histogram intersection value that determines similarity between the block in the first image and the block in the second image.

These and other aspects, objects, features and advantages of the present invention will be more clearly understood and appreciated from a review of the following detailed description of the preferred embodiments and appended claims, and by reference to the accompanying drawings.

### **ADVANTAGEOUS EFFECT OF THE INVENTION**

The present invention has the advantage of improved classification of images by utilizing block-based comparison that checks for similarity between two images, and for near-duplicate images.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

Fig. 1 is a block diagram illustrating an overview of the present invention;

Fig. 2 is a block diagram illustrating a date and time clustering technique of the present invention;

Fig. 3 is a graph illustrating a scaling function used to map the result of the 2-means clustering;

Fig. 4 is a graph illustrating a typical result of the scaling function of Fig. 3;

Fig. 5 is a diagram illustrating a block diagram of an event boundary checking after the date and time clustering;

5 Fig. 6 is a diagram illustrating grouping of images within each event based on content;

Fig. 7 is a block diagram of a group-merging step of the present invention;

10 Fig. 8 is a block diagram of image re-arrangement within each group;

Fig. 9 is a flowchart of block-based histogram correlation technique;

Fig. 10 is diagram illustrating the comparison between block histogram;

15 Fig. 11 is a diagram illustrating the extraction of the main subject;

Fig. 12 is diagram of an example of best intersection mapping for three segment analysis;

Fig. 13 is an illustration of shift detection within the block based histogram correlation; and

20 Fig. 14 is also an illustration of shift detection within the block based histogram correlation.

### **DETAILED DESCRIPTION OF THE INVENTION**

25 In the following description, the present invention will be described in the preferred embodiment as a software program. Those skilled in the art will readily recognize that the equivalent of such software may also be constructed in hardware.

30 Still further, as used herein, computer readable storage medium may comprise, for example; magnetic storage media such as a magnetic disk (such as a floppy disk) or magnetic tape; optical storage media such as an optical disc, optical tape, or machine readable bar code; solid state electronic storage devices

such as random access memory (RAM), or read only memory (ROM); or any other physical device or medium employed to store a computer program.

In addition, the term "event" is defined herein as a significant occurrence or happening as perceived by the subjective intent of the user of the  
5 image capture device.

Before describing the present invention, it facilitates understanding to note that the present invention is preferably utilized on any well-known computer system, such a personal computer. Consequently, the computer system will not be discussed in detail herein. It is also instructive to note that the images  
10 are either directly input into the computer system (for example by a digital camera) or digitized before input into the computer system (for example by scanning).

Referring to now Fig. 1, there is illustrated a flow diagram illustrating an overview of the present invention. Digitized images are input into  
15 the computer system where a software program of the present invention will classify them into distinct categories. The images will first be ranked S10 in chronological order by analyzing the time and date of capture of each image. The date and/or time of capture of each picture may be extracted, for example, from the encoded information on the film strip of the Advanced Photo System (APS)  
20 images, or from information available from some digital cameras. The representations of the images will then be placed S20 into one of a plurality of distinct events by a date and time clustering analysis that is described below. Within each event, the contents of the images are analyzed S20 for determining whether images closest in time to an adjacent event should be maintained in the  
25 event as defined by the clustering analysis, or the adjacent events merged together. After the images are defined into events, a further sub-classification (grouping) within each event is performed. In this regard, the images within each event will then be analyzed by content S30 for grouping images of similar content together, and then the date and time S30 for further refining the grouping.

30 The event segmentation S20 using the date and time information is by a k-means clustering technique, as will be described in detail below, which

groups the images into events or segments. A boundary check is then performed on the segments S20 for verifying that the boundary images should actually be grouped into the segment identified by the clustering, as will also be described below.

5                    These groups of images are then sent to a block-based histogram correlator S30 for analyzing the content. For each event or segment sent to the correlator, a content or subject grouping S30 is performed thereon for further sub-classifying the images by subject within the particular event segment. For example, within one event, several different subjects may appear, and these  
10                    subject groupings define these particular subjects. The subject grouping is based primarily on image content, which is performed by a block-based histogram correlation technique. This correlation compares portions of two images with each other, as will also be described in detail below. The result of the ranking is the classification of images of each segment into distinct subject groupings. The  
15                    date and time of all the images within each subject grouping are then compared to check whether any two or more subject grouping can be merged into a single subject grouping S30.

                    A refinement and subject re-arrangement analysis S40 will further improve the overall classification and the subject grouping by rearranging certain  
20                    images within a subject group.

                    Referring to Fig. 2, there is shown an exploded block diagram illustrating the data and time clustering technique S20. First, the time interval between adjacent pictures (time difference) is computed S20a. A histogram of the time differences is then computed S20b, an example of which is shown in block  
25                    10. The abscissa of the histogram is preferably the time in minutes, and the ordinate of the histogram is the number of pictures having the specified time difference. The histogram is then mapped S20c to a scaled histogram using a time difference scaling function, which is shown in Fig. 3. This mapping is to take the pictures with small time difference and substantially maintain its time difference,  
30                    and to take pictures with a large time difference and compress their time difference.

A 2-means clustering is then performed S20d on the mapped time-difference histogram for separating the mapped histogram 10 into two clusters based on the time difference; the dashed line represents the separation point for the two clusters. For further details of 2-means clustering, Introduction to Statistical  
5 Pattern Recognition, 2<sup>nd</sup> edition by Keinosuke Fukunaga 1990 may be consulted, and therefore, the process of 2-means clustering will not be discussed in detail herein. Referring briefly to Fig. 4, the result of 2-means clustering is the segmentation of the histogram into two portions 10a and 10b. Normally, events are separated by large time differences. The 2-means clustering, therefore, is to  
10 define where these large time differences actually exist. In this regard, the right portion 10b of the 2-means clustering output defines the large time differences that correspond to the event boundaries.

Referring to Fig. 5, there is illustrated an example of boundary checking between events. For two consecutive events i and j, a plurality of block-  
15 based, histogram comparisons are made to check if the pictures at the border of one event are different from the pictures at the border of the other event. If the comparison of content is similar, the two segments are merged into one segment. Otherwise, the segments are not merged. Preferably, the comparisons are performed on the three border images of each event (i3, i4, i5 with j1, j2, j3), as  
20 illustrated in the drawing. For example, image i5 is compared with image j1 and etc. This block-based histogram comparison technique will be described in detail hereinbelow.

Referring to Fig. 6, there is illustrated an overview of subject (content) grouping for each segmented event. Within each segmented event i,  
25 adjacent pictures are compared (as illustrated by the arrows) with each other using the below-described, block-based histogram technique. For example, the block-based histogram technique may produce five subject groupings (for example groups 1-5) from the one event i, as illustrated in the drawing. The arrangement of the subject grouping is stored for future retrieval during the subject arrangement  
30 step S40. In particular, the subject grouping having a single image is stored (for example, groups 2, 3, and 5).

Referring to Fig. 7, after the grouping by content, a time and date ordering is performed on the groupings for merging groups together based on a time and date analysis. A histogram of the time difference between adjacent images in the event is computed, similar to Fig. 4. A predetermined number of the

5 largest time differences (for example boundary  $a_{12}$ ) are compared with the boundaries (for example boundaries  $b_{12}$ ,  $b_{23}$ ,  $b_{34}$ ,  $b_{45}$ ) of the subject grouping determined by the block-based analysis. The predetermined number of largest time differences are determined by dividing the total number of images within an event by the average number of picture per group (four is used in the present

10 invention). If the boundary of the subject grouping matches the boundary based on the chosen time differences, the subject groupings will not be merged. If there is not a match between the two boundaries, the subject groupings having a boundary not having a matched time difference in the histogram will be merged into a single subject grouping (for example groups  $b_1$ ,  $b_2$ ,  $b_3$  merged into resulting

15 group  $c_1$ ).

Referring to Fig. 8, there is illustrated a diagram of image re-arrangement within each group. The arrangement of the initial subject groupings is retrieved for identifying subject groupings that contain single images (for example the groups with a single image of Fig. 6 –groups 2, 3, and 5 that are re-

20 illustrated as groups 2, 3, and 5 in Fig. 8). Any single images from the same subject grouping that are merged as identified by the merged subject grouping (for example, groups  $c_1$  and  $c_2$  of Fig. 7) are compared with all other images in the merged subject grouping, as illustrated by the arrows. This comparison is based on block-based histogram analysis. If the comparisons are similar, these images

25 will be re-arranged so that the similar images are located adjacent each other, for example groups  $d_1$  and  $d_2$ .

Further refinement is done by comparing any group that still contains a single image after the above procedure, with all the images in the event. This is to check if these single image groups can be better arranged within the

30 event grouping. This comparison is similar to the subject re-arrangement step of Fig. 8.

Referring to Fig. 9, there is illustrated a flowchart of the block-based histogram correlation used in the above analyses. First, a histogram of the entire image of both images is computed S50, a global histogram. A comparison of the two histograms is performed by histogram intersection value S60 illustrated the following equation:

$$Inter (R, C) = \frac{\sum_{i=1}^n \min (R_i, C_i)}{\sum_{i=1}^n R_i}, \quad \text{Eq. 1}$$

where R is the histogram of the reference image, C is the histogram of the candidate image, and n is the number of bins in the histogram. If the intersection is under a threshold S65, preferably 0.34, although other thresholds may be used, the images are different. If the threshold is met or exceeded S65, then a block-based histogram correlation will be performed S70. In this regard, each image will be divided into blocks of a given size, preferably 32 x 32 pixels in the present invention. It is instructive to note that those skilled in the art may vary the block size depending on the resolution of the image without departing from the scope of the invention. For each block, a color histogram is computed. Referring to Fig. 10, if one image is considered a reference image and one image a candidate image, the images are compared in the following way. Each block 20 of the reference image is compared to the corresponding block 30 of the candidate image and to the adjacent blocks 40, 8 blocks in the present invention.

Referring to Fig. 9, the block histograms between the reference image and the candidate image are compared using the histogram intersection equation defined above S80. The average intersection value is derived by computing the average of the best intersection values from each of the block comparisons S90. This average intersection value will be compared to a low threshold (preferably 0.355), and a high threshold (preferably 0.557). If the average intersection value is below the low threshold S95, the two images are considered different. If the average intersection value is above the high threshold S96, then the two images are considered similar. If the average intersection value

higher than the near duplicate threshold S97, preferably .75, then the images are considered near duplicates. It is instructive to note that near duplicates are significant because it permits the user to select the best of the two images to use in the albuming process.

- 5                    If the average intersection value is between these two thresholds, further analysis will be performed as described below (main subject/background analysis or the 3-segment average intersection analysis S100).

- 10                   Referring to Figs. 9 and 11, in some instances, it is desirable to see if the foreground or main subject 65 are similar between two images even though the backgrounds 66 are different. For example, there may be two pictures of your dog in two different places, but it will be desirable to group them together because they are the same subject. The images are divided into two parts, a background 66 and foreground 65 as shown in Fig. 11. The corresponding parts are then compared between the two images to see if there is any similarity between the two
- 15                   main subject areas by using the histogram intersection value of Equation 1. The average of this histogram intersection value from the block comparison will be compared with a threshold. If the average value is greater than the threshold, preferably .5, than the two main subjects are considered similar. It is instructive to note that any well-known algorithm may be used to extract the main subject area
- 20                   for comparison, in lieu of the above described method. Such alternative embodiments are within the intended scope of this invention.

- In an alternative embodiment, and referring to both Figs. 9 and 12, another method for performing 3-segment analysis to determine if the two images may contain a similar subject is shown. In this regard, this is performed by first
- 25                   forming a map 60 which contains the average of the two highest intersection values of each of the block comparisons; for example, 9 comparisons were performed in the illustration of Fig. 10, the average of the highest two will be used for map 60. Fig. 12 illustrates, for example, a 9 x 6 block although it should be understood that the map size depends on the size of the image. This map is
- 30                   divided into three parts: the left portion 70a, the center portion 70b, and the right portion 70c. If the average intersection value of the center portion 70b is higher

than a threshold (preferably 0.38) S105, the two images may contain a very similar subject in the center portion 70b of the image, and the two images may be considered to be similar by subject. In addition, the comparisons of the histogram will be performed with the reference and candidate images reversed. If the two  
5 images are similar both methods should give substantially similar correlation; obviously if they are different, the results will not be similar. The images are then checked S110 to determine if there is a high intersection value in one of the directions, right, left, up, and down.

Referring to Figs. 9, 13 and 14, shift detection is used to determine  
10 the case when the two images 90 and 100 have very similar subject that appears in different locations of the image. For example, the main subject may be situated in the center of one image and to the left-hand side of the other image. Such a shift can be determined by recording both the best intersection values of the reference blocks, as well as the coordinates of the corresponding candidate blocks. For each  
15 block of the reference image, nine block to block comparisons are made, the two best intersection values and the directions associated with the two best intersection values are used for determining the shift direction. The shift is performed on any combination of four basic directions, north, east, south and west. There are two cases which should be noted for clarity: either the directions are opposite (example  
20 N and S) or adjacent (example N and W). If the two directions are adjacent, a one-block shift (of the overall candidate image) in the combined direction, for example one north and one east gives a shift in the NE direction, as illustrated in Fig. 13. On the other hand if the directions are opposite, two one-block shifts, one in each direction, are performed, as illustrated in Fig. 14 The above analysis and the shift  
25 can be repeated S120 to check for similarity.

The invention has been described with reference to a preferred embodiment. However, it will be appreciated that variations and modifications can be effected by a person of ordinary skill in the art without departing from the scope of the invention.

**CLAIMS:**

1. A method for comparing image content of first and second images, the method comprising the steps of:
  - (a) extracting a portion of both the first and second images both of which portions are determined to include a main subject area of each image;
  - (b) dividing the main subject area of the images into a plurality of blocks;
  - (c) computing a color histogram for one block in each image;
  - (d) computing a histogram intersection value between the block of the first image and the block of the second image; and,
  - (e) determining a first threshold value for the computed histogram intersection value that determines similarity between the block in the first image and the block in the second image.
2. The method as in claim 1 further comprising shifting the first image in a predetermined direction if there is not similarity between the first and second images.
3. The method as in claim 1, wherein step (c) includes (c1) computing a color histogram for each of the plurality of blocks, and step (d) includes (d1) computing a histogram intersection value between one block in the first image and a plurality of blocks in the second image.
4. The method as in claim 3 further comprising shifting the first image in a predetermined direction if there is not similarity between the first and second images.
5. The method as in claim 4 further comprising repeating step (d1) if there is not similarity between the first and second images.

6. The method as in claim 1, wherein step (a) includes dividing both of the images into three portions and using a middle portion of the three portions as the main subject area.

7. The method as in claim 1 further comprising the step of determining a second threshold value for the computed histogram intersection value that determines similarity between the block in the first image and the block in the second image and consequently if the two images are near-duplicates of each other.

8. A computer program product for comparing image content of first and second images comprising: a computer readable storage medium having a computer program stored thereon for performing the steps of:

(a) extracting a portion of both the first and second images both of which portions are determined to include a main subject of each image;

(b) dividing the main subject of the images into a plurality of blocks;

(c) computing a color histogram for one block in each image;

(d) computing a histogram intersection value between the block of the first image and the block of the second image; and,

(e) determining a threshold value for the computed histogram intersection value that determines similarity between the block in the first image and the block in the second image.

9. The computer program product as in claim 8 further comprising the step of shifting the first image in a predetermined direction if there is not similarity between the first and second images.

10. The computer program product as in claim 8, wherein step (c) includes (c1) computing a color histogram for each of the plurality of blocks,

and step (d) includes (d1) computing a histogram intersection value between one block in the first image and a plurality of blocks in the second image.

11. The computer program product as in claim 9 further comprising shifting the first image in a predetermined direction if there is not similarity between the first and second images.

12. The computer program product as in claim 11 further comprising repeating step (d1) if there is not similarity between the first and second images.

13. The computer program product as in claim 12, wherein step (a) includes dividing both of the images into three portions and using a middle portion of the three portions as the main subject area.

14. The computer program product as in claim 8 further comprising the step of determining a second threshold value for the computed histogram intersection value that determines similarity between the block in the first image and the block in the second image and consequently if the two images are near-duplicates of each other.

## ABSTRACT

A method for comparing image content of first and second images, the method comprises the steps of extracting a portion of both the first and second images both of which portions are determined to include a main subject area of each image; dividing the main subject area of the images into a plurality of blocks; computing a color histogram for one block in each image; computing a histogram intersection value between the block of the first image and the block of the second image; and determining a first threshold value for the computed histogram intersection value that determines similarity between the block in the first image and the block in the second image.

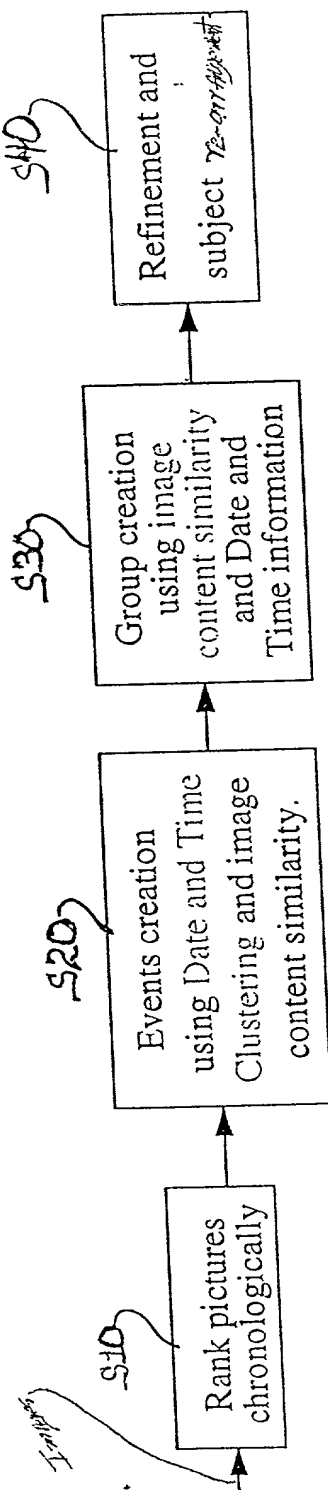


Fig. 1

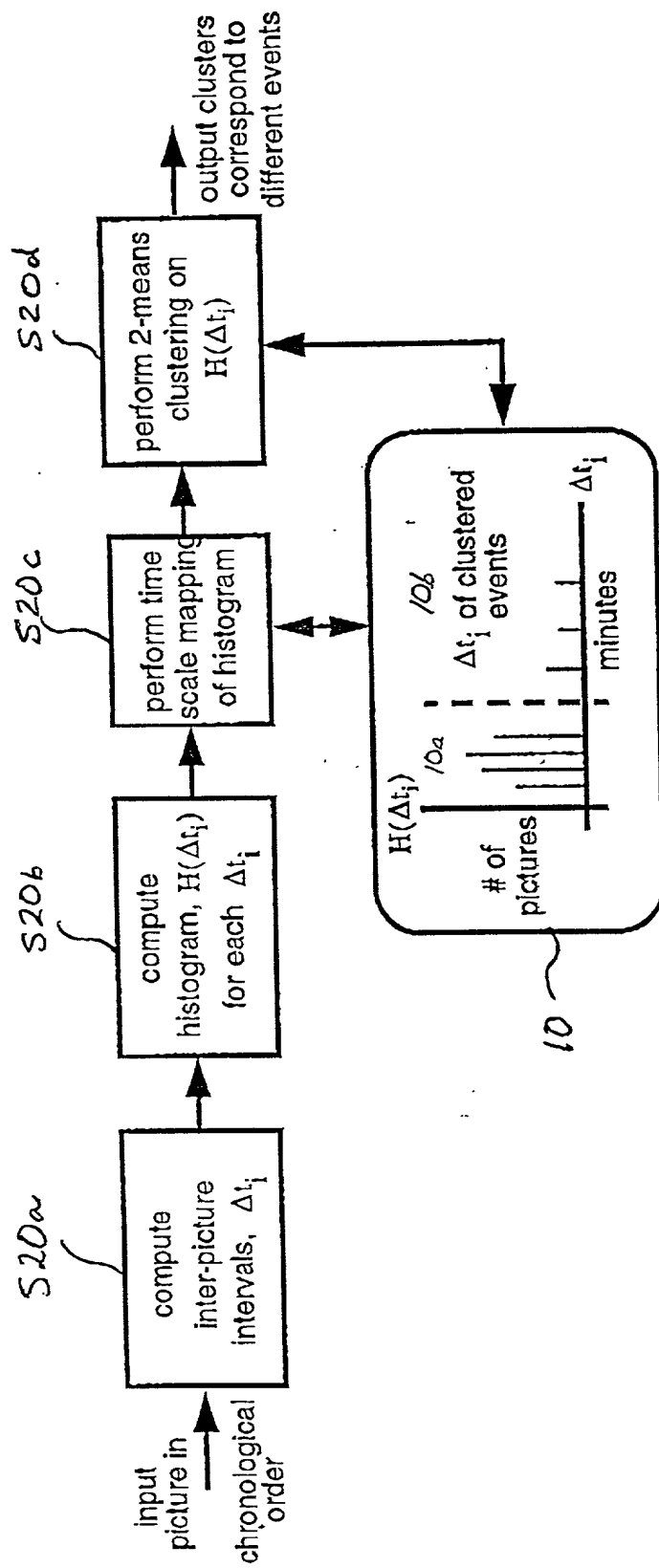


Fig. 2

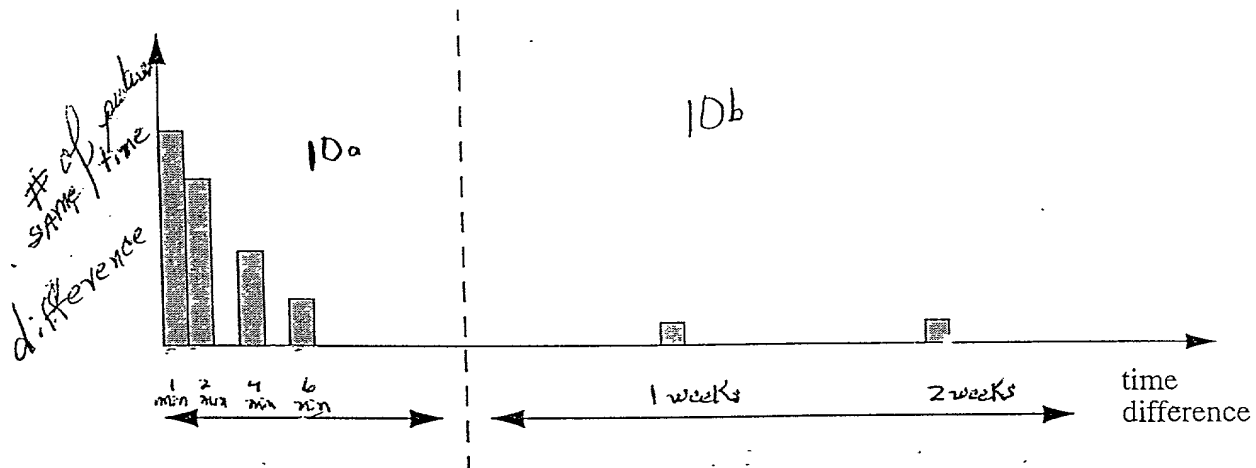


Fig. 4

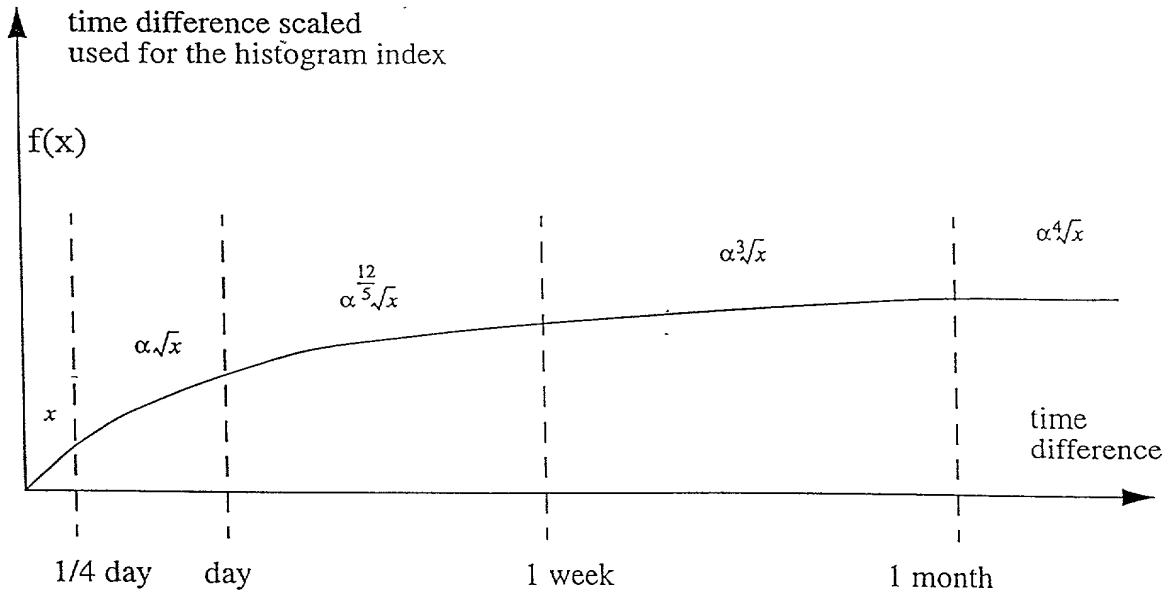


Fig. 3

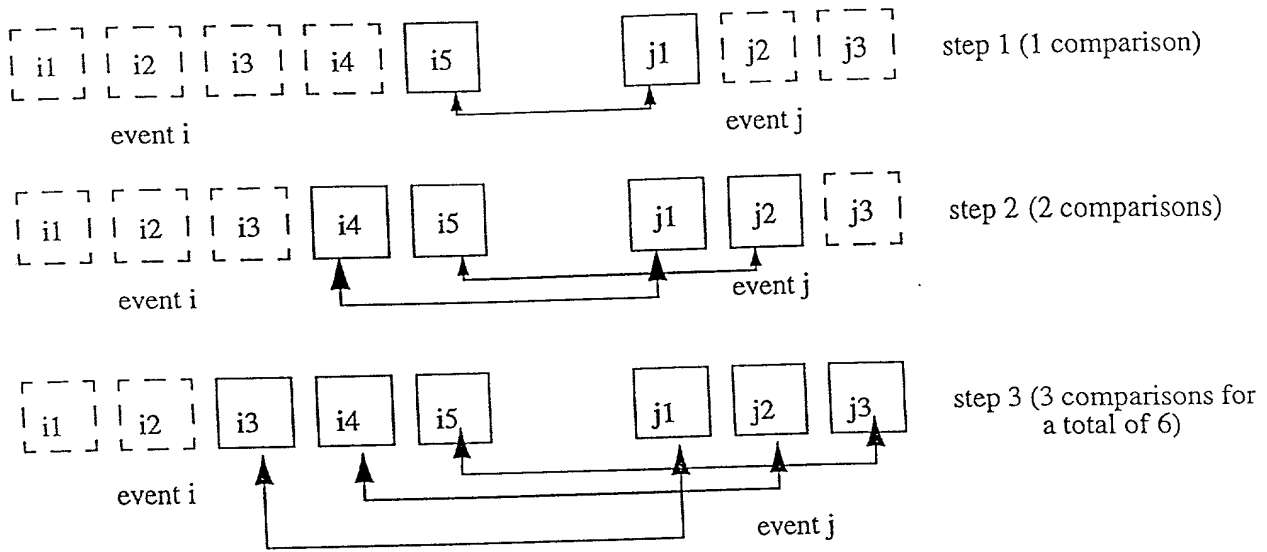


Fig. 5

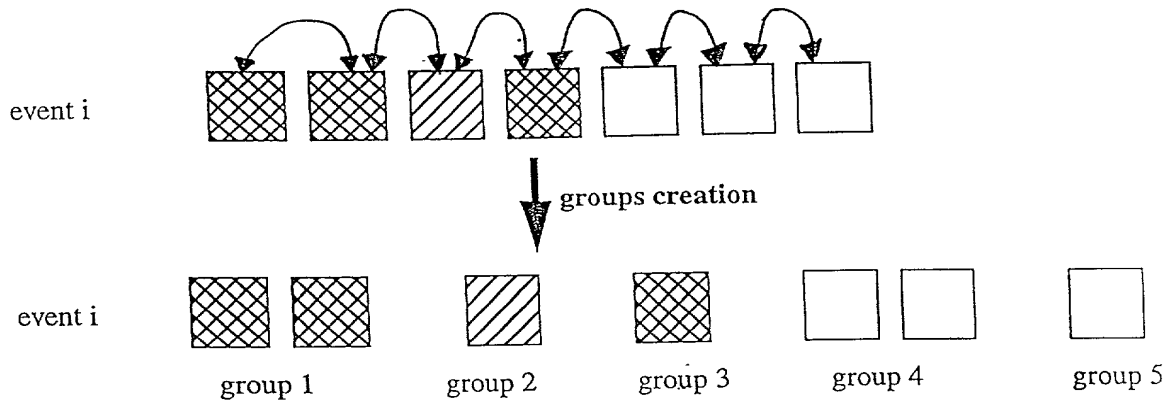


Fig. 6

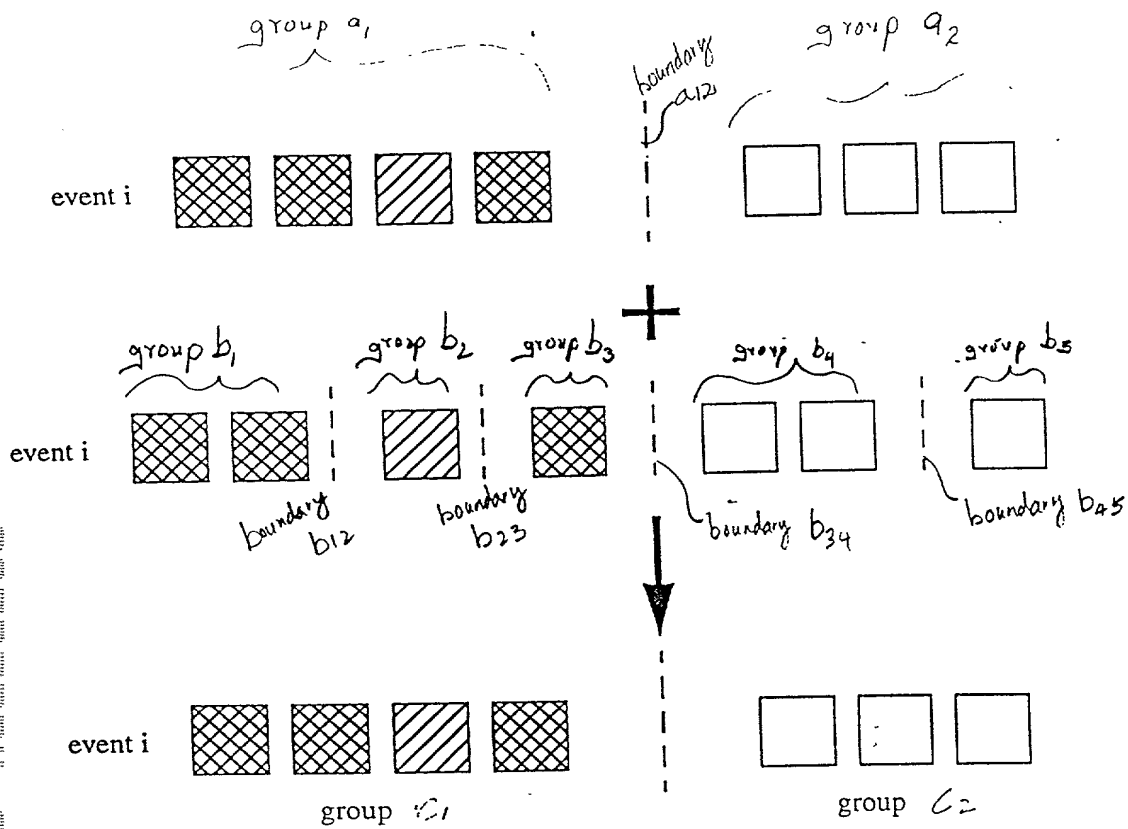


Fig. 7

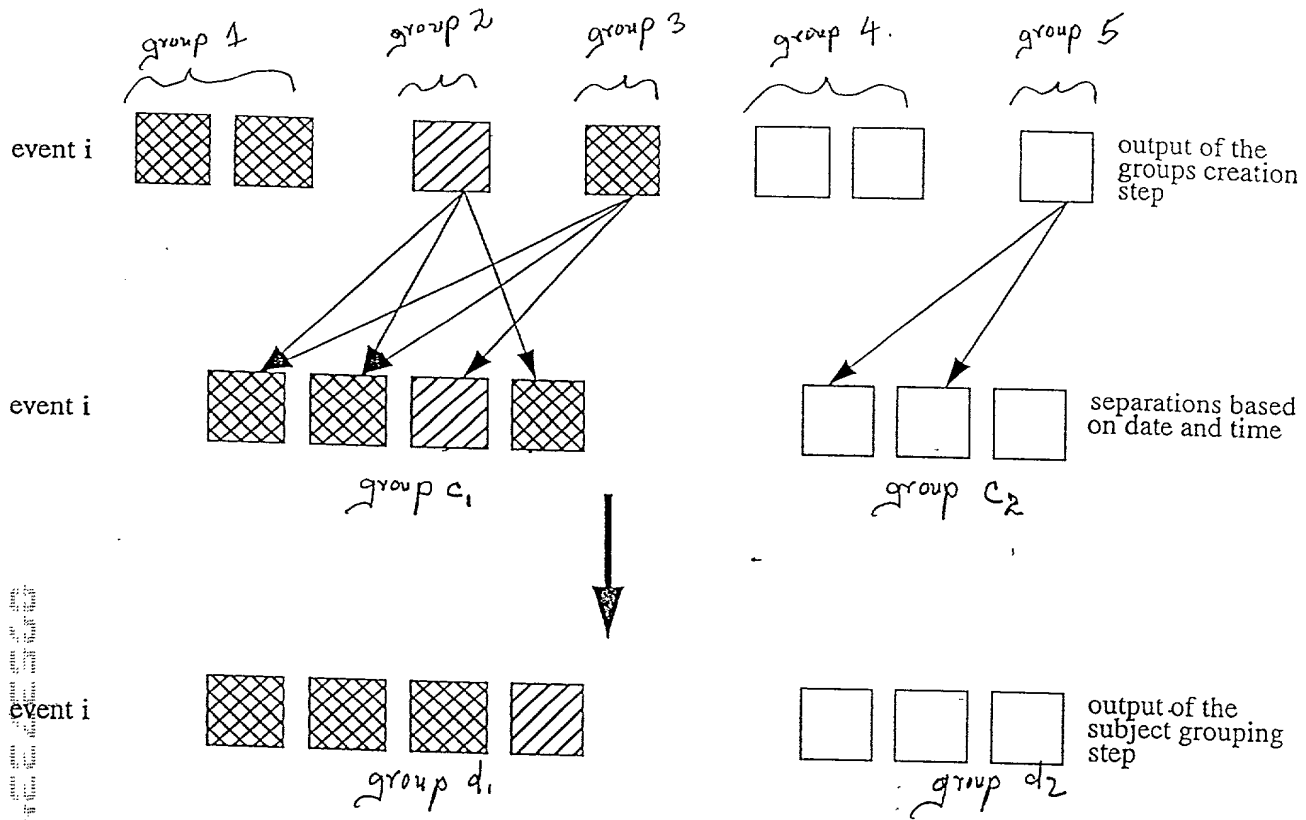


Fig. 8

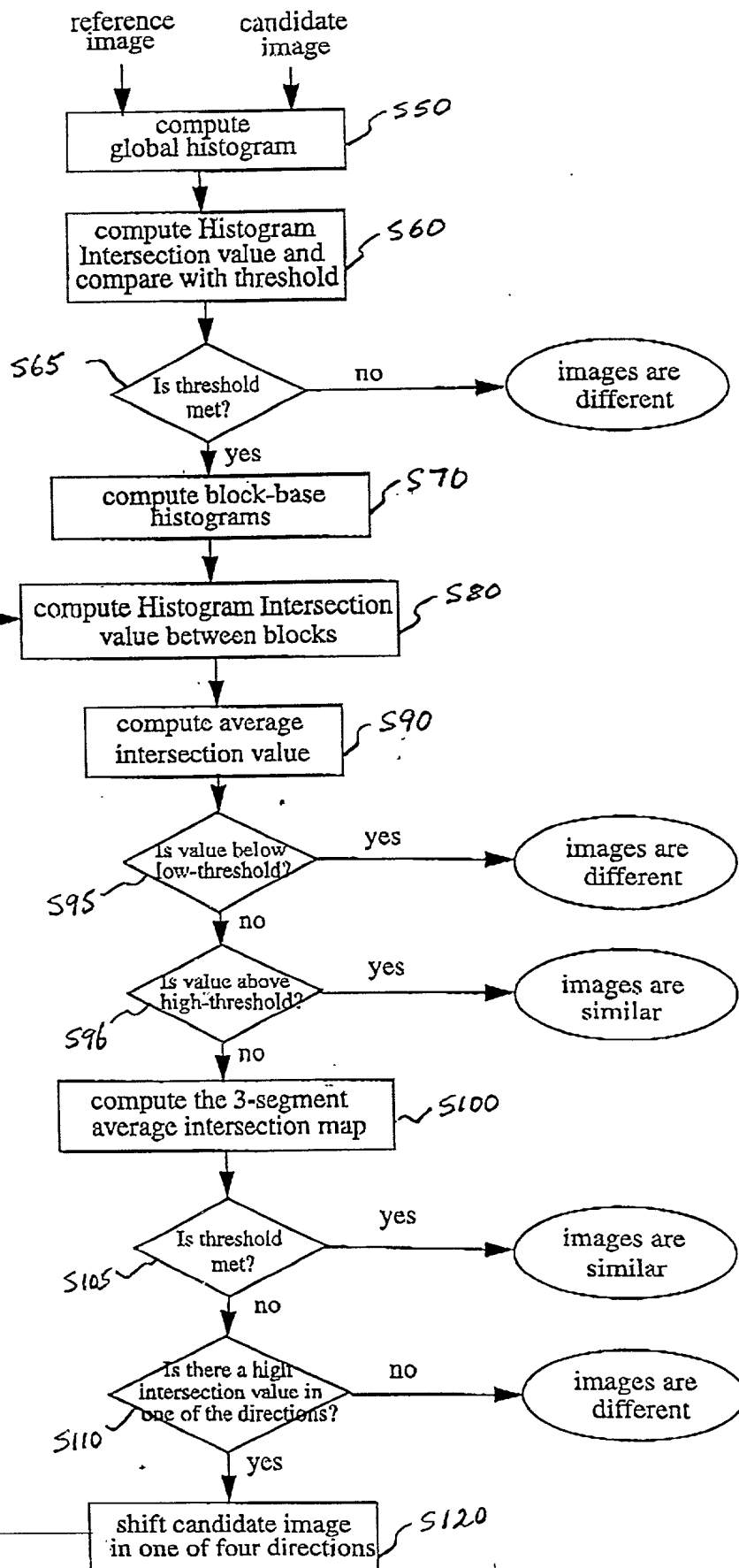


Fig. 9

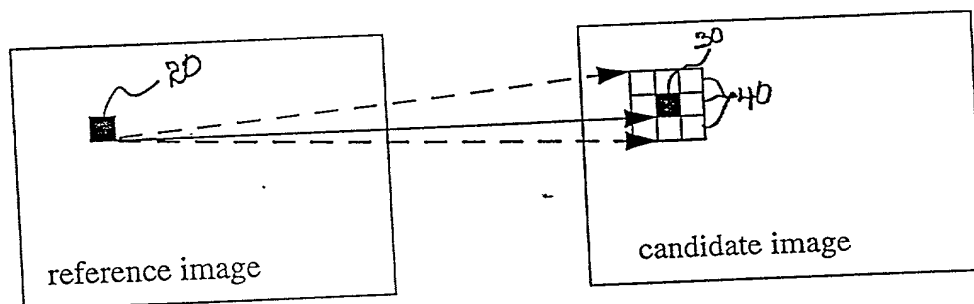


Fig. 10

$70a$        $\frac{60}{70b}$        $70c$

0.32	0.45	0.11	0.34	0.56	0.76	0.87	0.24	0.09
0.22	0.10	0.07	0.41	0.48	0.80	0.77	0.44	0.26
0.03	0.16	0.37	0.46	0.68	0.69	0.75	0.21	0.02
0.21	0.34	0.24	0.56	0.87	0.51	0.48	0.11	0.14
0.26	0.22	0.19	0.27	0.38	0.23	0.31	0.12	0.25
0.01	0.08	0.18	0.15	0.18	0.21	0.12	0.20	0.21

Fig. 11

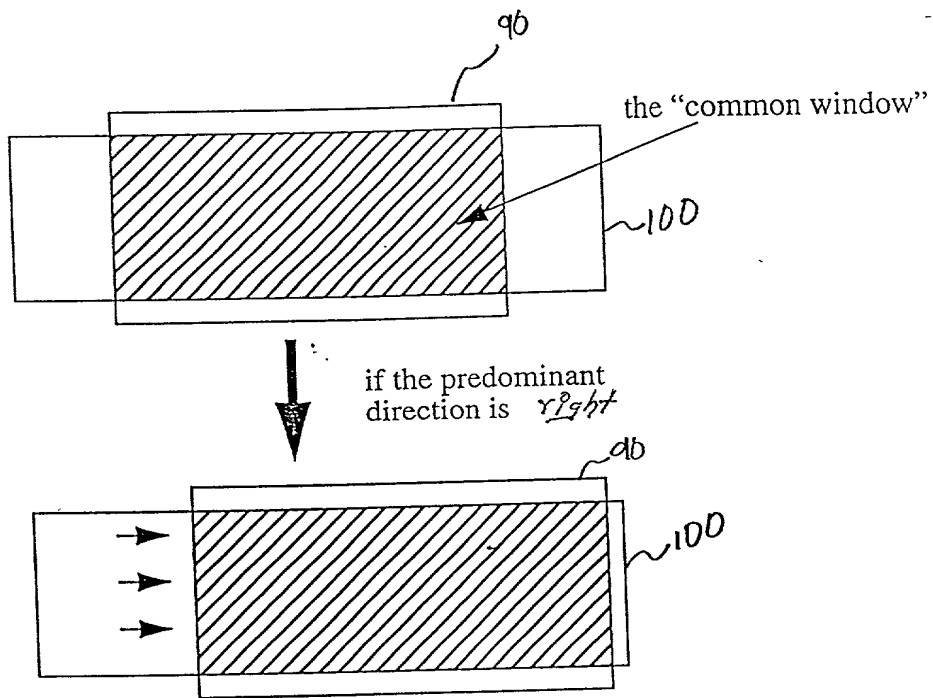


Fig. 12

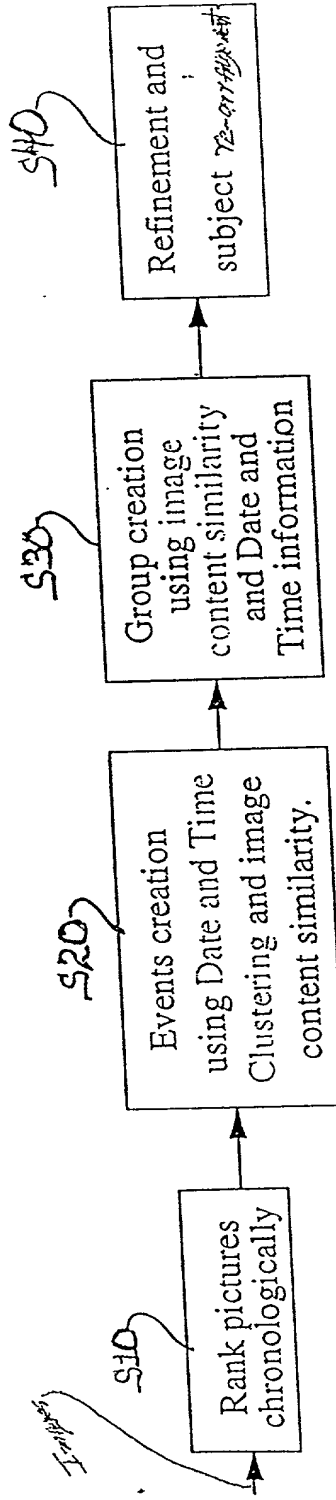


Fig. 1

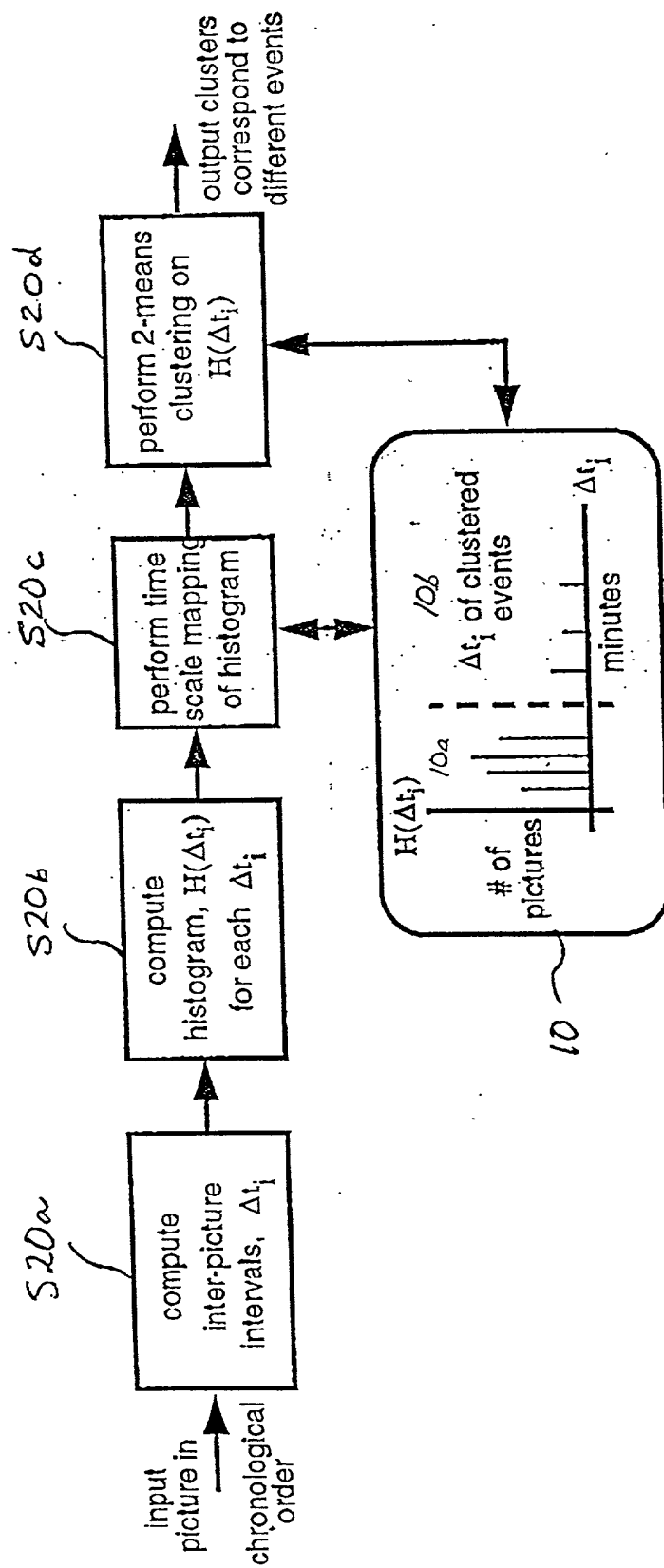


Fig. 2

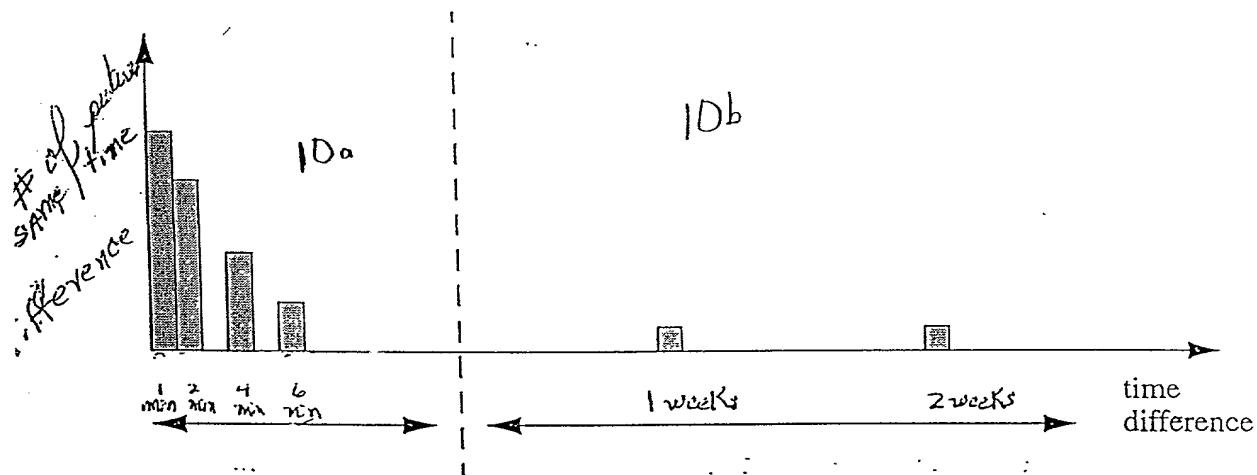


Fig. 4

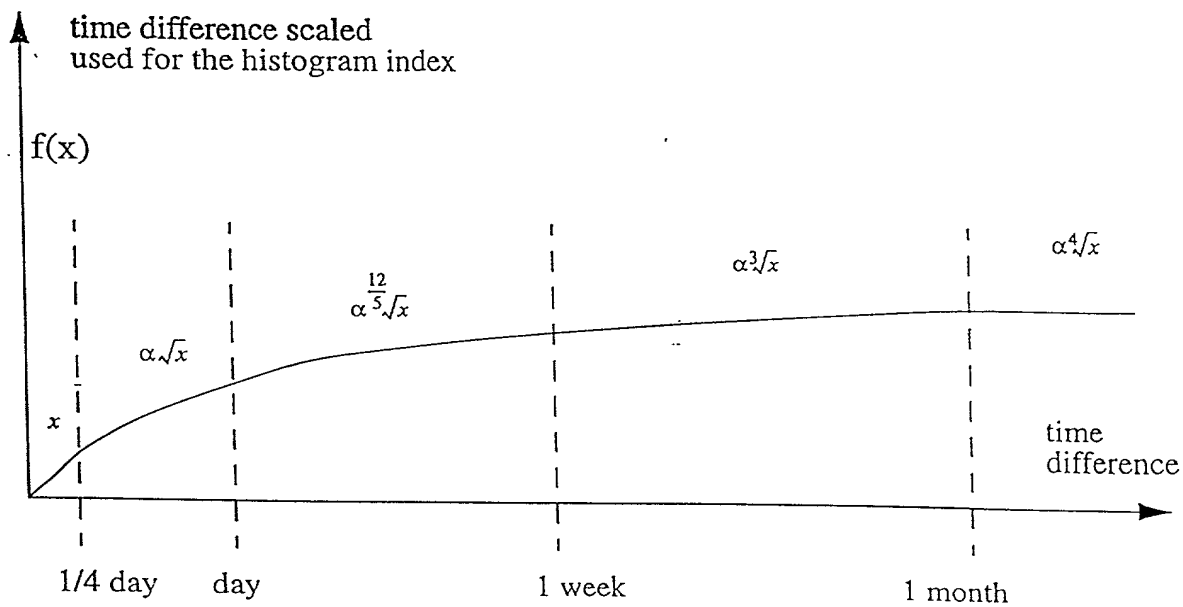


Fig. 3

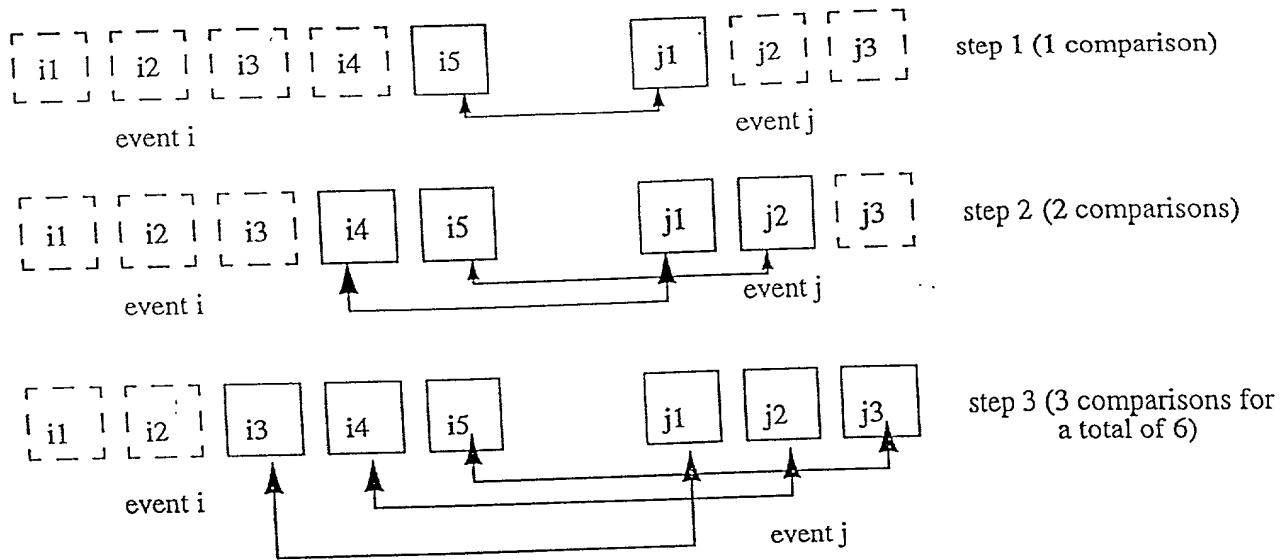


Fig. 5

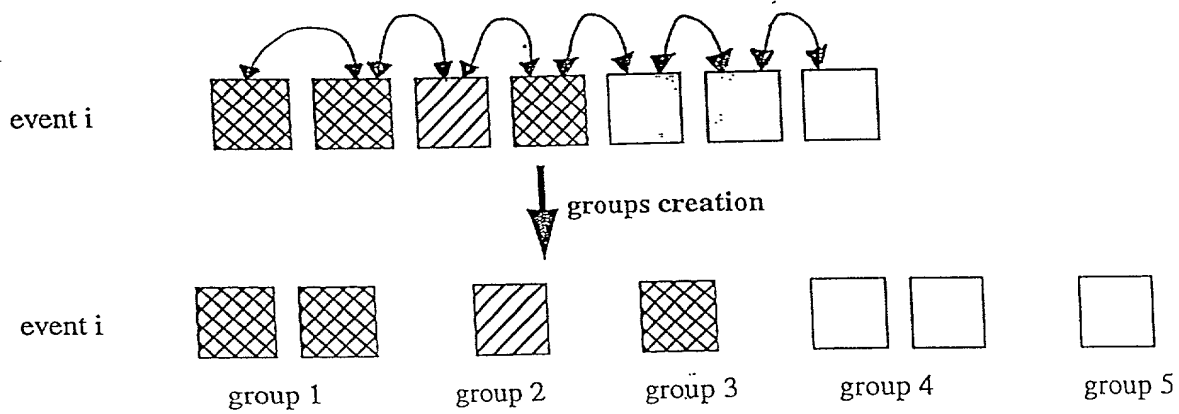


Fig. 6

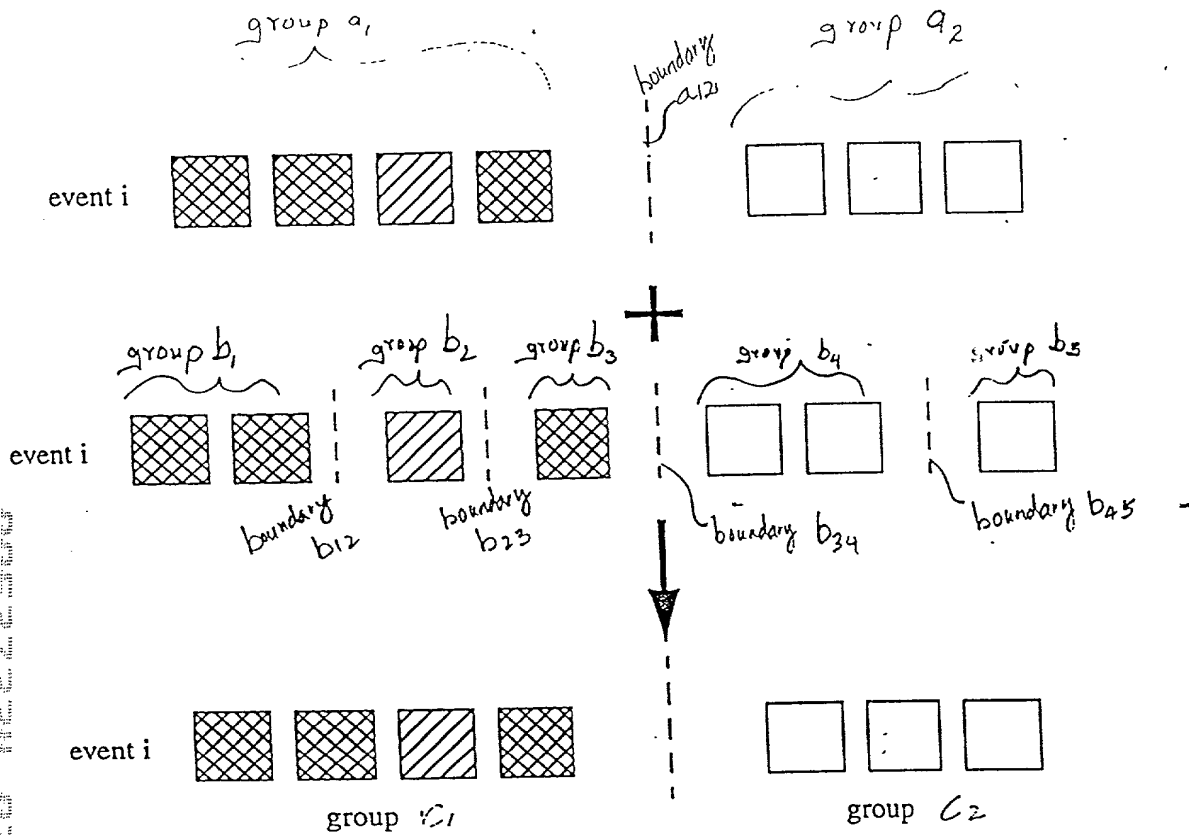


Fig. 7

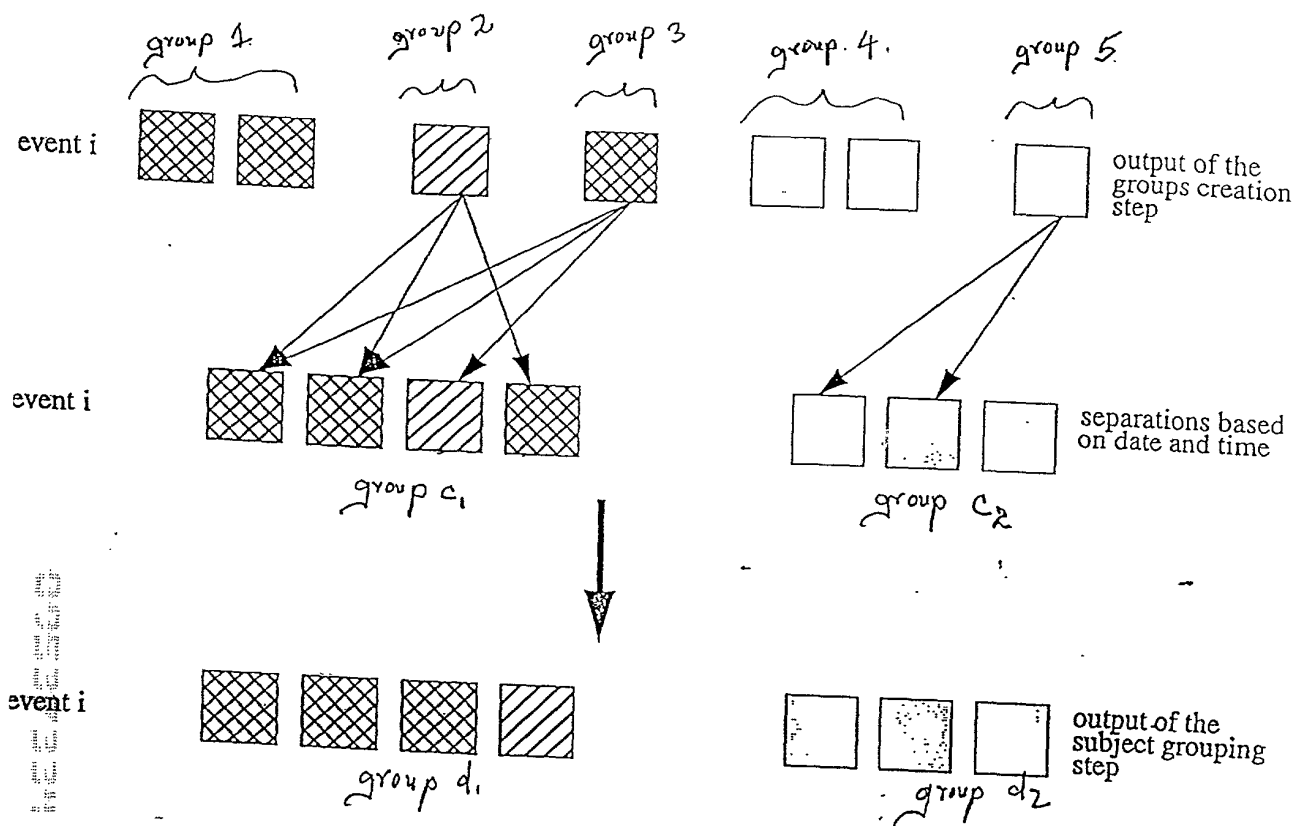


Fig. 8

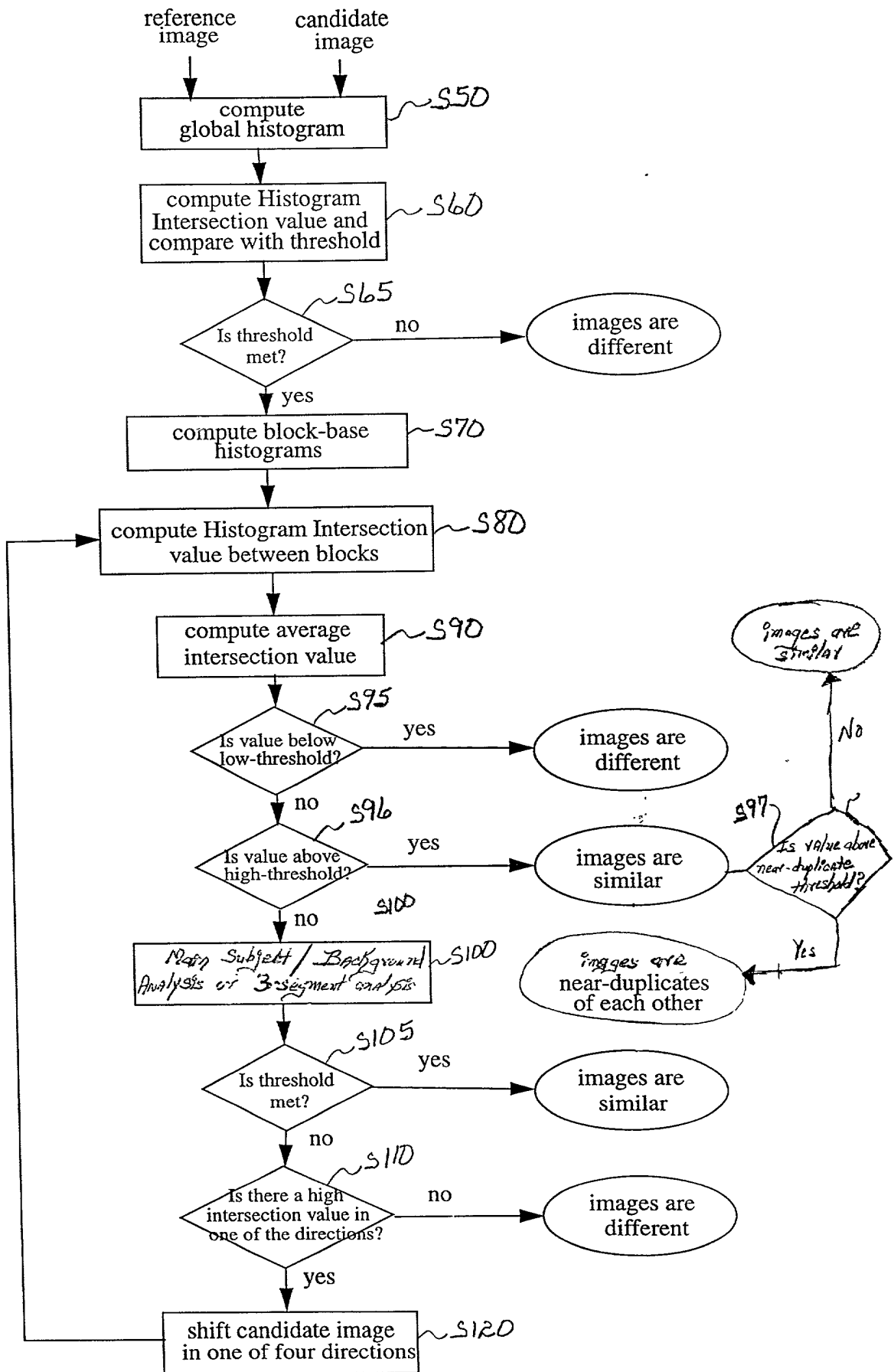


Fig. 9

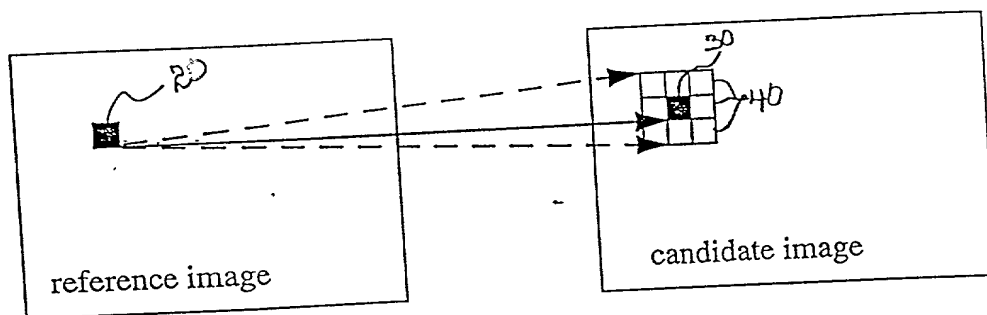


Fig. 10

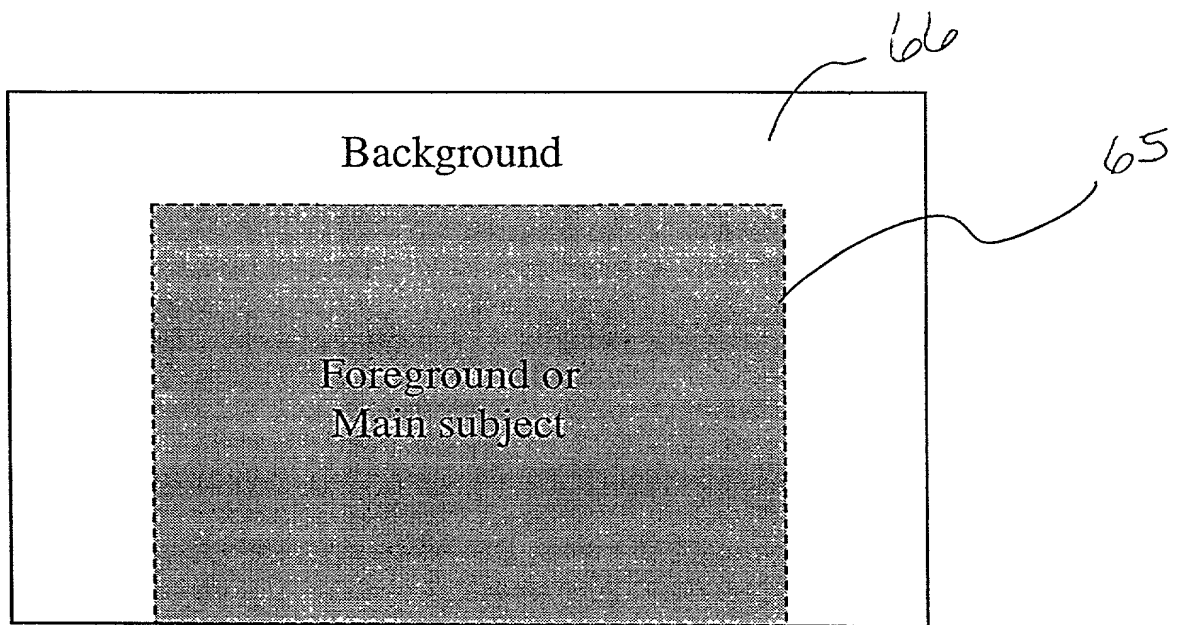


Fig. 11

70a
60  
70b
70c

0.32	0.45	0.11	0.34	0.56	0.76	0.87	0.24	0.09
0.22	0.10	0.07	0.41	0.48	0.80	0.77	0.44	0.26
0.03	0.16	0.37	0.46	0.68	0.69	0.75	0.21	0.02
0.21	0.34	0.24	0.56	0.87	0.51	0.48	0.11	0.14
0.26	0.22	0.19	0.27	0.38	0.23	0.31	0.12	0.25
0.01	0.08	0.18	0.15	0.18	0.21	0.12	0.20	0.21

Fig. 12

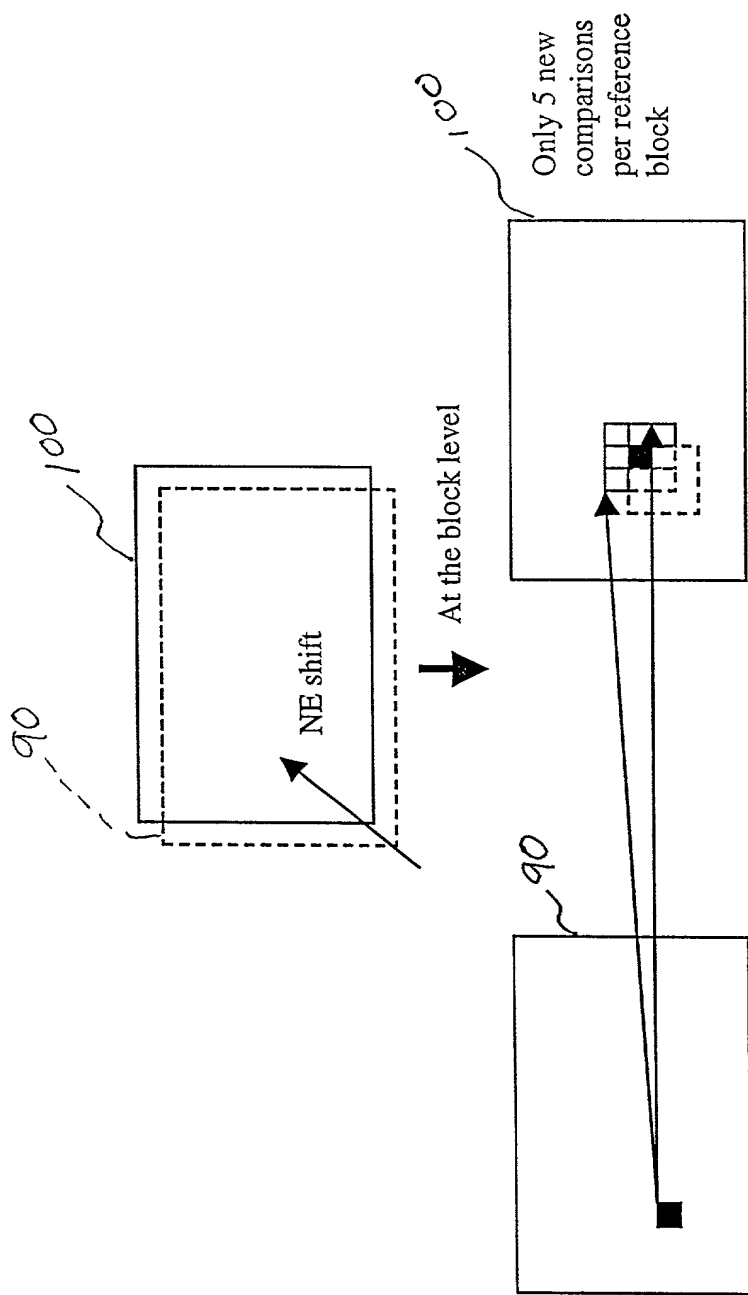


Fig. 13

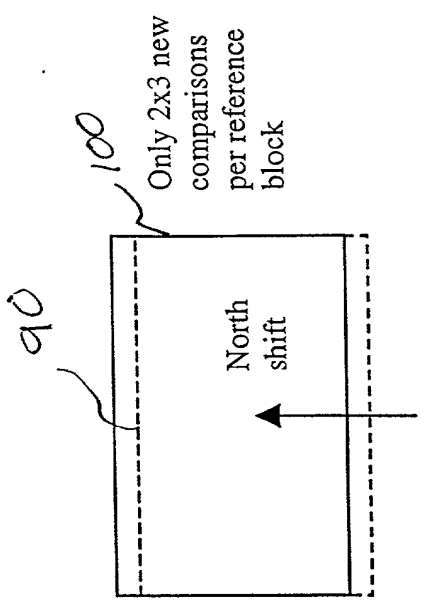
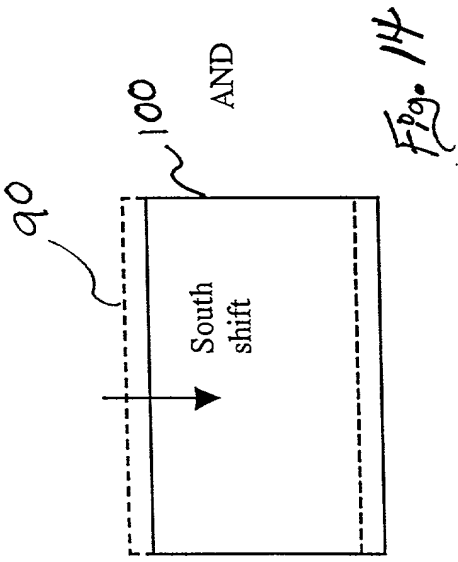


Fig. 14

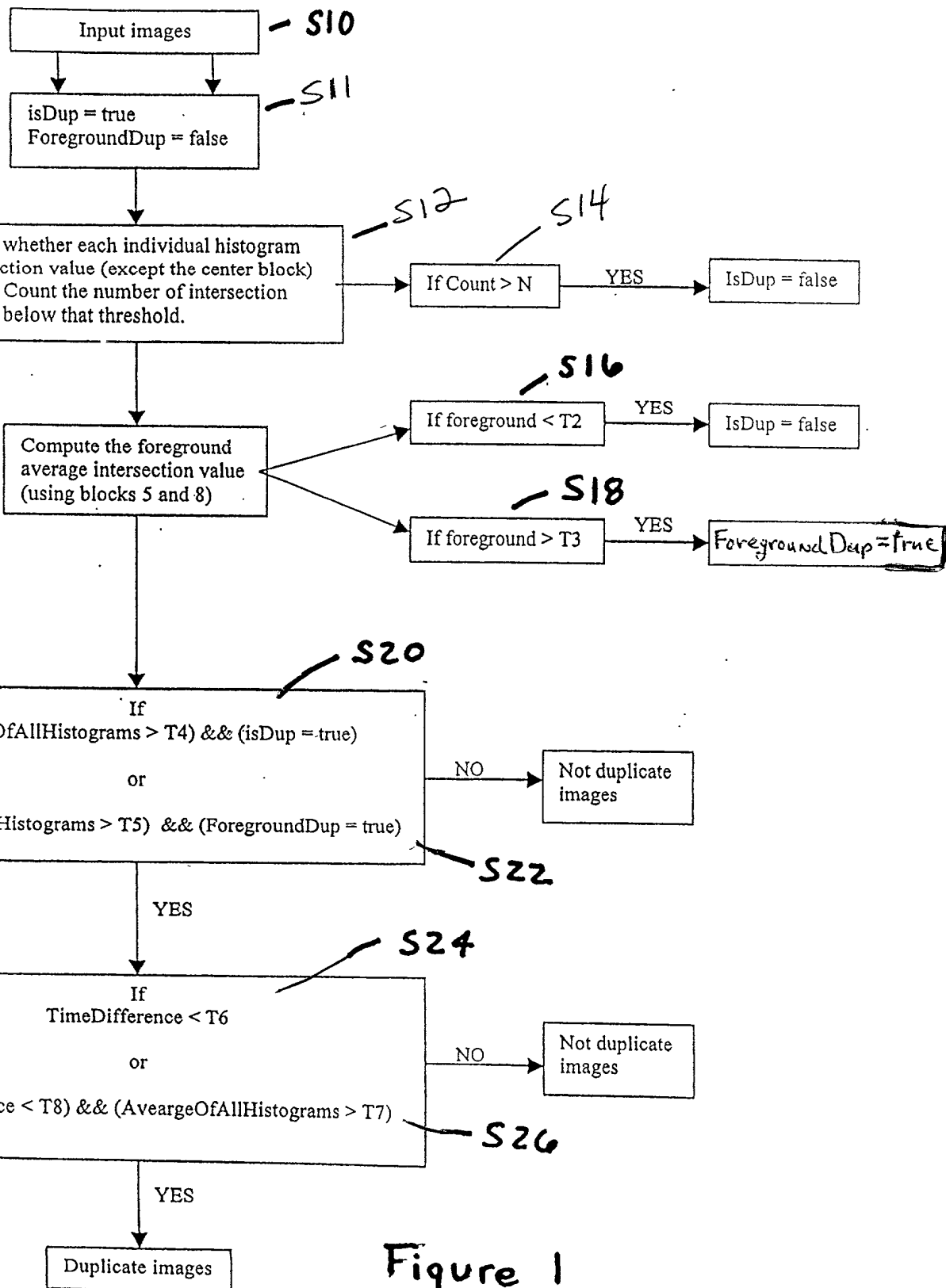
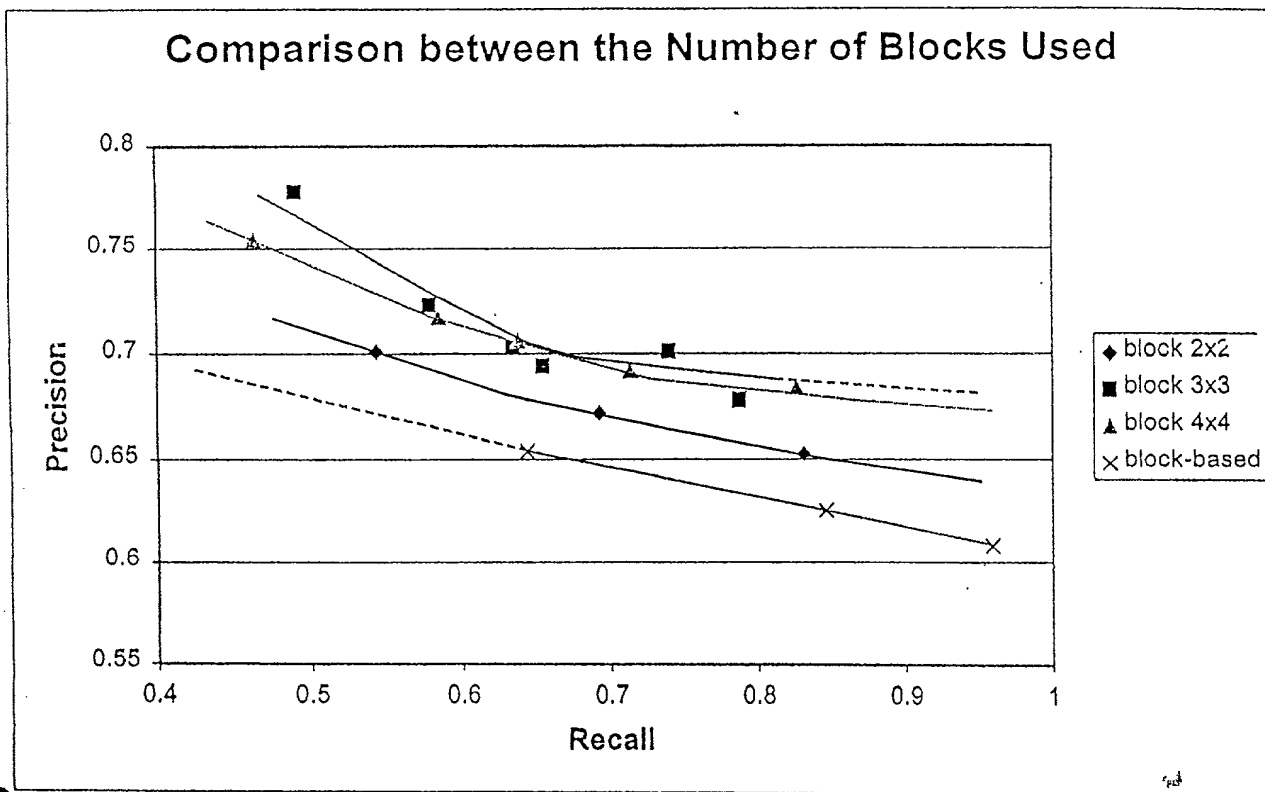
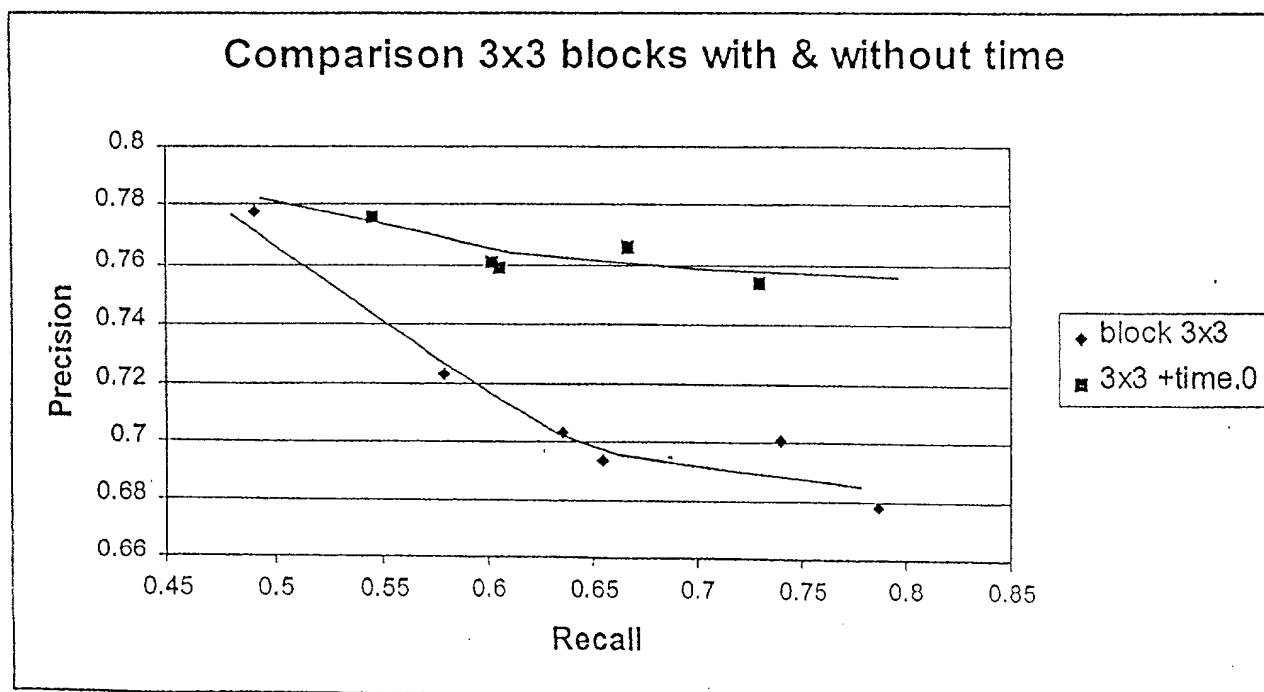


Figure 1



**Figure 2** Comparison between the number of blocks used.



**Figure 3** Comparison 3x3 blocks with and without time

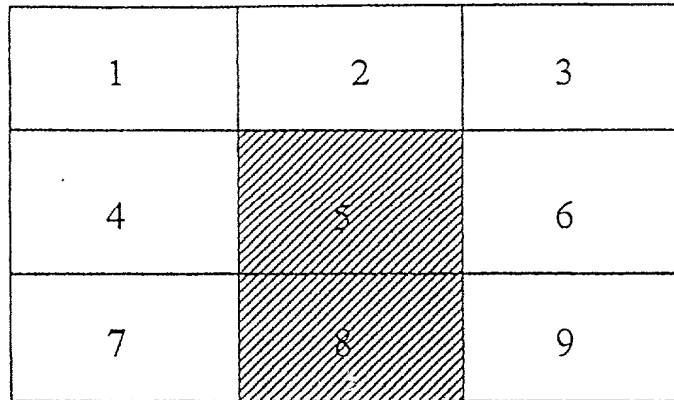


Fig. 4 Representation of blocks 5 and 8 in an image

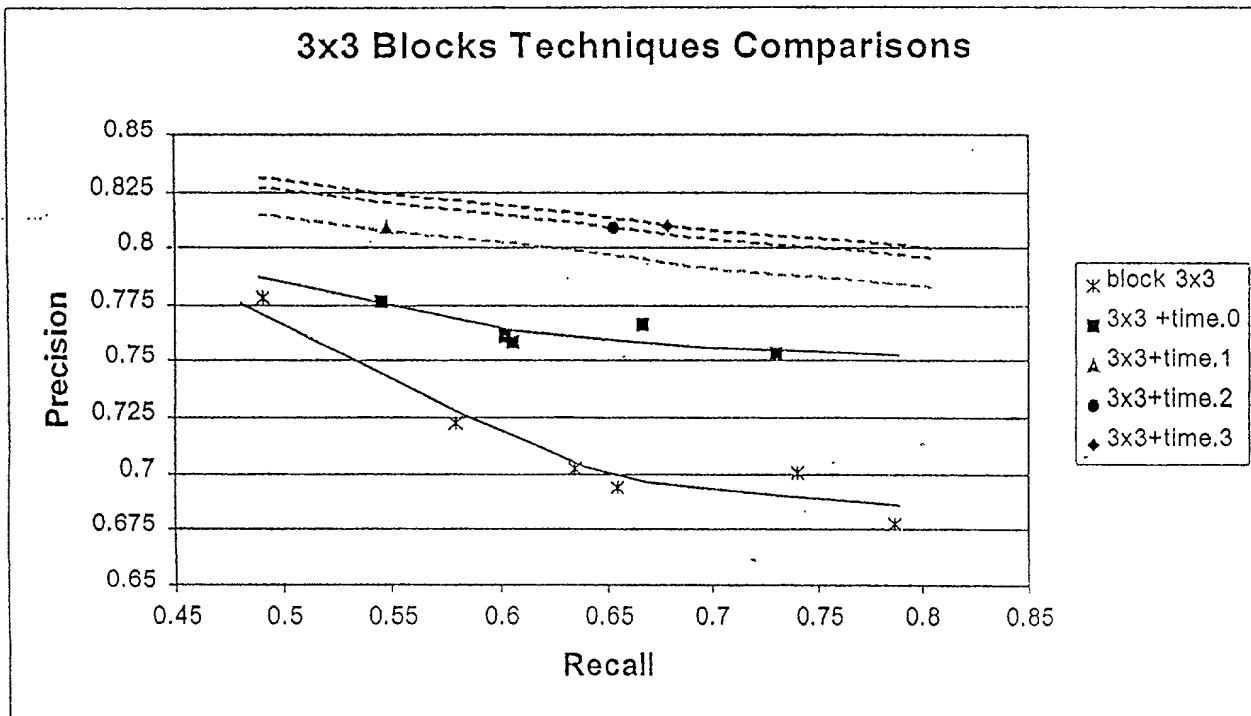
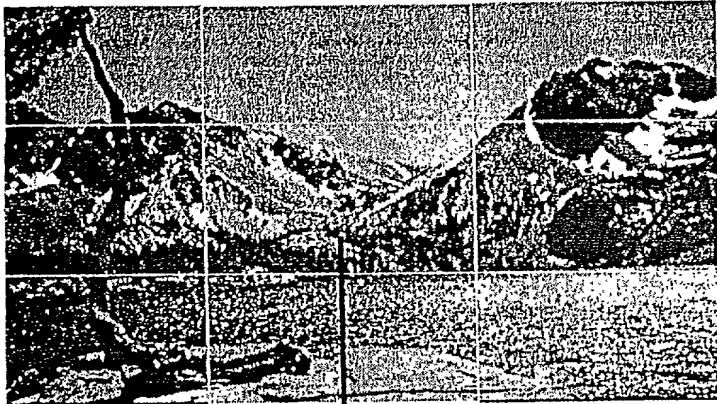


Figure 6 3x3 Blocks Techniques Comparison

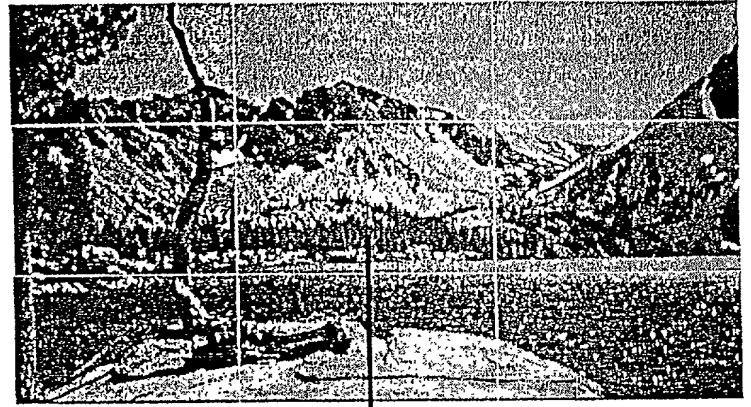
Block-based histogram technique	3x3 blocks technique	3x3 blocks technique with time information	3x3 blocks technique with time final version
R = 0.644	R = 0.655 (+1.7%)	R = 0.667 (+1.8%)	R = 0.679 (+1.9%)
Pr = 0.654	Pr = 0.694 (+6.1%)	Pr = 0.766 (+10.4%)	Pr = 0.810 (+5.7%)

Figure 7

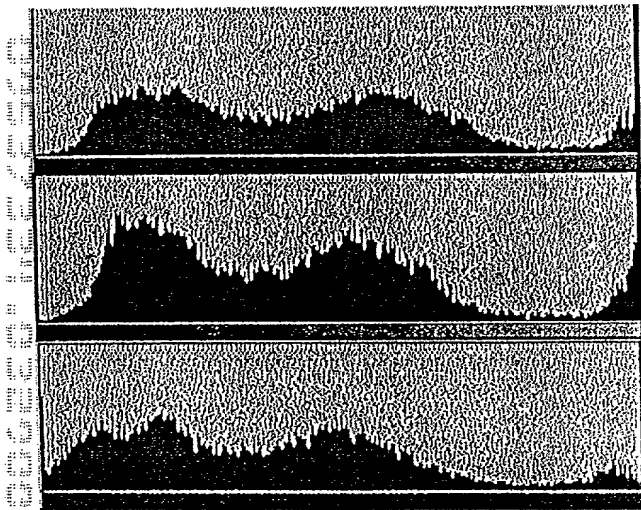
Figure 5 gives an illustration of how the blocks of each image are compared to each other using color histograms.



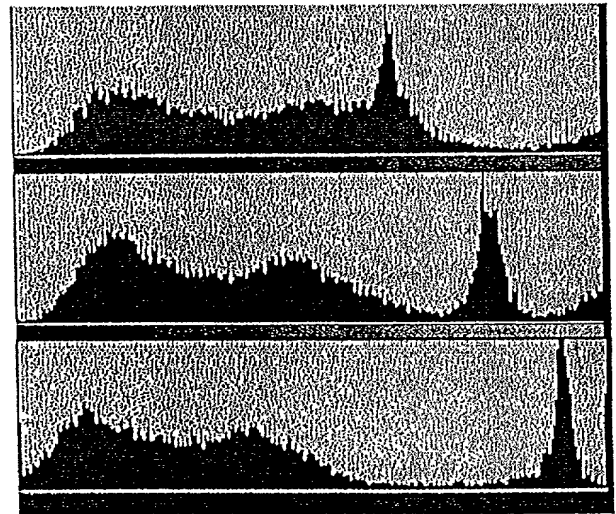
Picture A



Picture B



Color histograms (R, G, B) of the center block of Picture A



Color histograms (R, G, B) of the center block of Picture B

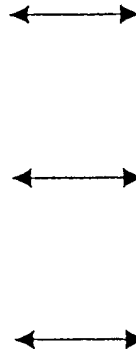


FIG. 5A

FIG. 5B

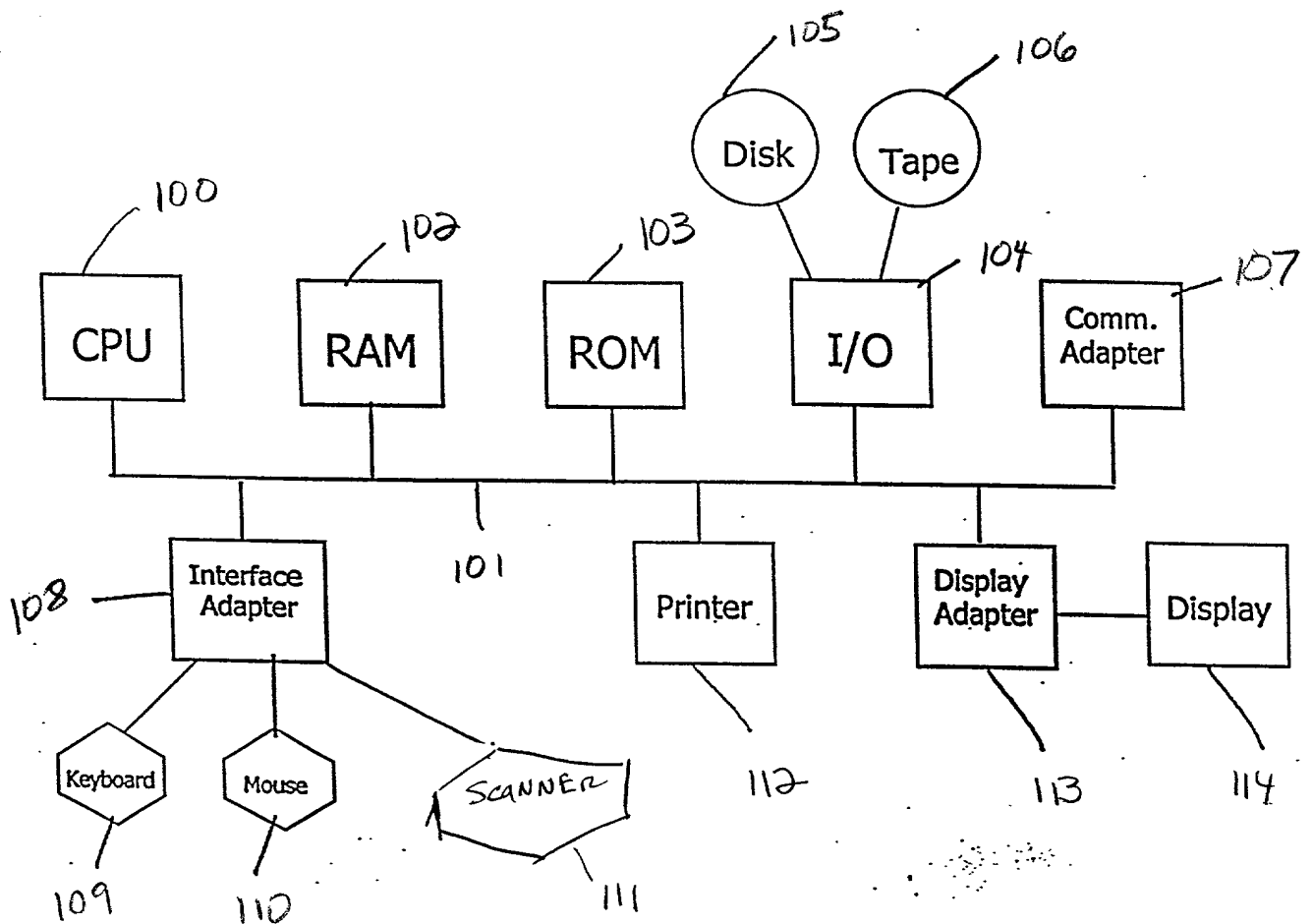


FIG. 8

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Alexander C. Loui  
Eric S. PavieA METHOD OF DETECTING  
DUPLICATE PICTURES IN AN  
AUTOMATIC ALBUMING SYSTEM

Serial No. To Be Assigned

Filed 29 March 2000

Commissioner for Patents  
Washington, D.C. 20231

Sir:

Group Art Unit:

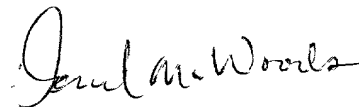
Examiner:

I hereby certify that this correspondence is being  
deposited today with the United States Postal Service  
as first class mail in an envelope addressed to  
Commissioner for Patents, Washington, D.C. 20231.  
Robin G. Reeves3/29/00  
DateLETTER UNDER RULE 53

Pursuant to Rule 53, the above-identified application, enclosed herewith (including specification and claims), is being filed without a signed declaration or assignment in the names of the inventors, Alexander C. Loui, Eric S. Pavie. The declaration and assignment will be filed later.

Please address all correspondence to Thomas H. Close, Patent Legal Staff, Eastman Kodak Company, Rochester, New York 14650-2201. Please direct all telephone communications to David M. Woods at (716) 477-5256.

Respectfully submitted,

David M. Woods  
Attorney for Applicant(s)  
Registration No. 27,171DMW/RGR  
Telephone: (716) 477-5256  
Facsimile: (716) 477-4646